Title
The roles of E-Tmod in mechanotransduction and cardiac myofibrillogenesis

Permalink
https://escholarship.org/uc/item/44w7918z

Author
Lian, Ian Yu-Zen

Publication Date
2007

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA, SAN DIEGO

The Roles of E-Tmod in Mechanotransduction and Cardiac Myofibrillogenesis

A dissertation submitted in partial satisfaction of the requirements for the Degree Doctor of Philosophy in Bioengineering

by

Ian Yu-Zen Lian

Committee in charge:

Professor Shu Chien, Chair
Professor L. Amy Sung, Co-Chair
Professor Yu-hwa Lo
Professor Jeff Omens
Professor Jason Yuan

2007
The dissertation of Ian Yu-Zen Lian is approved, and it is acceptable in quality and form for publication:

________________________________________

________________________________________

________________________________________

________________________________________

Co-Chair

Chair

University of California, San Diego

2007
## TABLE OF CONTENTS

| Signature Page | .......................................................... | iii |
| Table of Contents | .......................................................... | iv |
| List of Figures | ........................................................... | vi |
| List of Tables | .......................................................... | xi |
| Acknowledgements | .......................................................... | xii |
| Vita and publications | .......................................................... | xiii |
| Abstract | .......................................................... | xv |

### CHAPTER I

**Introduction**

A. The building blocks of the cardiac contractile machinery ................ 1  
B. Components of the cardiac sarcomere ........................................ 3  
C. The cardiac myofibril Z-disk ................................................... 4  
D. Overview of the sarcomeric defects and cardiac phenotypes ............ 4  
  D.1. Phenotypes of sarcomere mutations ...................................... 4  
  D.2. Phenotypes of sarcomere deletions ...................................... 6  
E. Introduction to the Tropomodulin (Tmod) gene family .................... 7  
F. Significance of Erythrocyte Tropomodulin (E-Tmod) in the heart ...... 10  
G. General Objective and approach ............................................ 10  

### CHAPTER II

**Gene expression profiling of the E-Tmod Deficient Mouse** ............. 12  
Abstract .................................................................................. 12  
Introduction ............................................................................. 12  
Material and Methods .................................................................. 14  
  A. Sectioning and staining of whole heart tissues .......................... 14  
  B. Strategy for high-throughput expression profiling of *E-Tmod*<sup>−/−</sup> 
Cardiac tissue ........................................................................... 15  
  C. Collection of mouse hearts ...................................................... 16  
  D. Preparations of cardiac RNA and cDNA .................................. 16  
  E. Mouse oligonucleotide microarray ........................................ 17  
  F. Microarray pre-hybridization and hybridization ....................... 17  
  G. Scanning and image quantification of microarrays .................... 18  
  H. Designation of flag (low quality) spots .................................. 18  
  I. Data mining and analysis ..................................................... 18  
  J. Defining differentially expressed (DE) genes ............................ 19  
  K. Quantitative real-time PCR reactions .................................... 20
Results .......................................................................................................................... 21
A. Cardiac phenotype of $E$-$Tmod^{+/−}$ and $E$-$Tmod^{−/−}$ mice……….. 21
B. E-Tmod expressions in $E$-$Tmod^{+/−}$ and $E$-$Tmod^{−/−}$ embryonic hearts……………………………………………………………………………………………………… 23
C. Expressions of Tmod homologs in $E$-$Tmod^{+/−}$ and $E$-$Tmod^{−/−}$ embryonic hearts……………………………………………………………………………………………………… 24
D. Number of differentially expressed genes in $E$-$Tmod^{+/−}$ cardiac tissues……………………………………………………………………………………………………… 25
E. Clusters of differentially expressed genes in $E$-$Tmod^{+/−}$ cardiac tissues……………………………………………………………………………………………………… 26
F. Clusters of differentially expressed genes analyzed by GeneSet Enrichment Analysis (GSEA)……………………………………………………………………………………………………… 29
G. Validation of the microarray reported DE genes with qPCR assay………………………………………………………………………………………………………………………………… 31
H. Differentially expressed genes in $E$-$Tmod^{+/−}$ and $E$-$Tmod^{−/−}$ cardiac tissues……………………………………………………………………………………………………… 32
Discussion………………………………………………………………………………………… 32

CHAPTER III
Hypoxia-induced Pressure Overload of E-Tmod Deficient Mouse……… 37
Abstract………………………………………………………………………………………… 38
Introduction…………………………………………………………………………………… 39
Material and Methods………………………………………………………………………… 39
A. Hyposia conditioning of $E$-$Tmod^{+/−}$ mice………………………………………… 39
B. Collection of hearts from hypoxia treated mice………………………………………… 39
C. Functional analysis………………………………………………………………………… 40
D. Sectioning and staining of whole heart tissues………………………………………… 40
Results
A. Normalized heart weight before and after hypoxia conditioning……….. 41
B. Hematocrit response of $E$-$Tmod^{+/−}$ mice to hypoxia treatment……….. 42
C. Cardiac response to hypoxia…………………………………………………………….. 42
D. Functional analysis………………………………………………………………………. 44
Discussion………………………………………………………………………………………… 44

CHAPTER IV
Mechanical Regulation of E-Tmod in Cardiomyocytes………………… 47
Abstract………………………………………………………………………………………… 47
Introduction…………………………………………………………………………………… 47
Material and Methods………………………………………………………………………… 49
A. Cloning of E-Tmod cDNA from rat cardiomyocytes…………………………. 49
B. Cell culture…………………………………………………………………………………… 49
C. E-Tmod knockdown with siRNA…………………………………………………………. 50
D. Application of laminar shear…………………………………………………………….. 51
E. Application of mechanical stretch………………………………………………………. 51
Results
A. Identification of E-Tmod Splice Variants in NRC………………………………….. 52
B. Comparison of E-Tmod sequence in human, mouse and rat………………….. 53
C. Hypertrophic response of NRT under strain……………………………………….. 54
D. Hypertrophic response of NRT under shear……………………………………….. 55
E. Up-regulation of E-Tmod by shear and static stretch…………………………….. 55
F. Up-regulation of Actin and TM by shear and static stretch……… 56
G. Up-regulation of E-Tmod and ANP by static stretch after RNAi knockdown………………………………………….……... 57
Discussion 57

CHAPTER V
Regulation of E-Tmod During Cardiac Myofibrillogenesis…… 61

Abstract……………………………………………………………….. 61
Introduction……………………………………………………………. 61
Material and Methods …………………………………………..……. 63
A. Extraction of embryonic mouse hearts……………………… 63
B. Culture and seeding of mouse embryonic hearts…………….. 63
C. Isolation of total RNA and reverse transcription to cDNA…… 63
D. Isolation of total protein and separation of cellular fractions… 64
E. Protein assay and western blot…………………………………. 65
F. Quantitative real-time PCR…………………………………..… 65
G. Immunostaining………………………………………………… 66
Results……………………………………………………………..… 67
A. Localization of E-Tmod in developing cardiac sarcomere…… 67
B. Localization of two E-Tmod splice variants in mouse hearts… 68
C. Temporal expression of E-Tmod throughout cardiogenesis…… 68
D. Temporal expressions of TPM1 and ACTC throughout cardiogenesis………………………………………………………... 69
E. Correlation between temporal expressions of E-Tmod, TPM1,
and ACTC……………………………………………………………... 70
F. Correlation between thin-filament genes and HR during cardiogenesis……………………………………………………………... 71
G. Correlation between the expressions of thin-filament genes
and Fcsa……………………………………………………………. 72
H. Correlation between temporal expressions of thin filament
genes and peak ventricular pressure………………………….…….. 75
I. Expressions of α and β TPM during embryonic heart
development………………………………………………………….. 78
J. Expressions of ACTC and TPM1 in E-Tmod−/− embryonic
heart…………………………………………………………………… 78
K. Expressions of cardiac troponins in E-Tmod−/− embryos…………………………………………………………………… 79
L. Cardiac contraction in wildtype and E-Tmod−/− cultured
Embryonic hearts…………………………………………………….. 79
Discussions…………………………………………………………… 81
Future Work

Appendix

Development of the Data Processing Algorithms and an Integrated
Microarray Analysis Software Package…………………………… 89
Objective……………………………………………………………… 89
Introduction………………………………………………………….. 89
Development of microarray analysis pipeline………………………… 91
A. Image acquisition and spot quality control…………………… 93
B. Raw data treatment……………………………………………… 94
LIST OF FIGURES

Figure 1. Schematic representation of the sarcomeric multi-protein Complex…………………………………………………………………………… 3
Figure 2. Localization of E-Tmod in cardiac myofibrils during Cardiogenesis………………………………………………………………………… 10
Figure 3. Strategy of large-scale gene expression profiling scheme utilizing the combination of microarray and qPCR platforms… 15
Figure 4. Flowchart of the application of improved-algorithms to process microarray data……………………………………………………………………………………………………………………………… 19
Figure 5. Morphology of the E-Tmod embryo proper, whole heart, and cardiomyocyte staining at gestational age E9……………… 22
Figure 6. RV cavity and free wall volumes of E-Tmod and wildtype Mice……………………………………………………………………………………… 23
Figure 7. Expression of E-Tmod in E-Tmod−/−, E-Tmod+/+ and E-Tmod++/+ embryo hearts………………………………………………………………………… 24
Figure 8. Tmod homolog expression profile in E-Tmod, E-Tmod and E-Tmod E10 embryo propers……………………………………………………………………………………………………………………………… 25
Figure 9. Venn diagrams summarizing the number of differentially expressed genes ………………………………………………………………………………………………………………………………………………………….. 26
Figure 10. Percentage distribution of the differentially expressed genes passing three quantitative criteria………………………………………………………………………………………………………………………… 27
Figure 11. Percentage distribution of the down-regulated and up-regulated gene-clusters revealed by non-parametric GeneSet Enrichment Analysis (GSEA)……………………………………………………………………………………………………………………………………………… 30
Figure 12. ECG capture and synchronization with video of titanium dioxide labeled mouse heart……………………………………………………………………………………………………………………………………… 40
Figure 13. Normalized heart weight of mice without and with hypoxia treatment………………………………………………………………………………………………………………………………………………………… 41
Figure 14. Hematocrit levels of E-Tmod+/− and E-Tmod++/+ mice under normoxia and after the hypoxia treatment…………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………
Figure 26. Cardiac actin (ACTC) and tropomyosin (TPM) expressions in sheared and static-stretched cardiomyocytes.

Figure 27. E-Tmod expression in control and siRNA knockdown samples.

Figure 28. Identification and localization of two E-Tmod splice variants, E0-9 and E0/3-9.

Figure 29. Myofibril width (thickness). An increase in average myofibril thickness was observed during the mouse cardiac development.

Figure 30. Detection and localization of E-Tmod in developing mouse hearts.

Figure 31. Temporal expression of E-Tmod in mouse embryonic hearts measured by qPCR assay.

Figure 32. Temporal expressions of (A) TPM1 and (B) ACTC in mouse embryonic hearts measured by qPCR assay.

Figure 33. Correlation between E-Tmod (E2) with TPM1 and ACTC expressions.

Figure 34. Correlation between E-Tmod (E6) with TPM1 and ACTC expressions.

Figure 35. In vivo heart rate of the developing mouse embryos.

Figure 36. Temporal heart rate and E-Tmod expressions during Embryogenesis.

Figure 37. Temporal correlation between TPM1, ACTC expressions and the in vivo embryonic mouse heart rate.

Figure 38. Temporal expressions of (A) E-Tmod (E-2), (B) TPM1 and (C) ACTC vs. isometric force (Fcsa).

Figure 39. Temporal expressions of (A) E-Tmod (E), (B) TPM1 and (C) ACTC vs. peak systolic ventricular pressure.

Figure 40. Transition of α- and β-TM during myofibrillogenesis.

Figure 41. Expressions of (A) TPM1 and (B) ACTC in wildtype and E-Tmod embryonic hearts.

Figure 42. Expression of cardiac troponin genes in E-Tmod−/− and E-Tmod−/+ embryo propers.

Figure 43. Whole heart organ culture of E-Tmod−/− and E-Tmod−/+ samples.

Figure 44. Contraction rate of cultured embryonic hearts extracted from E-Tmod−/− and E-Tmod−/+ embryo samples.

Figure A1. Incorporate high and low scanning settings by normalizing to the common genes.

Figure A2. Flowchart of the microarray data processing pipeline.

Figure A3. Background signal profile of Cy3 channel being smoothed by the weighted moving average algorithm.

Figure A4. Linear and log distribution of single channel microarray.

Figure A5. Global normalization between channels in log scale to match median.

Figure A6. Treatment vs. control plots of raw and globally normalized microarray data.

Figure A7. Illustration of a weighted linear least squares regression fitting procedure.

Figure A8. Total vs. ratio intensity (M-A) plot of the microarray signals.
Figure A9.  Comparison of ratio distributions of 6 replicated microarray experiments………………………………………………… 109
Fig. A10. Illustration of cross array normalization scheme…………….. 111
Fig. A11. Definition of differentially-expressed genes according to ratio distribution pattern……………………………………… 116
Fig. A12. Venn diagram summarizing logics of microarray filt……….. 117
ACKNOWLEDGEMENTS

The completion of this thesis would not have been possible without the guidance of my committee members, professors Shu Chien, L. Amy Sung, Jeffery H. Omens, Yu-hwa Lo and Jason Yuan. The nature of my thesis project required a wide range of experimental resources from the laboratories of my committee members and many collaborators. Specifically, I would like to thank Kurt, Brian, and Eric from Professor Wolfgang Dillmann’s laboratory and Zhuangjie from Dr. McCulloch’s group.

I also want to express my gratitude to my doctoral colleagues for offering me their knowledge and skills. Among many, Drs. Carlos Vera, Weijuan Yao, Darrell Belke, Julie Li, Jane Li, Jason Haga, Troy Hornberger, and Bruce Jacobson contributed significantly to my research and helped me through many troubleshooting sessions.

During my studies in the graduate program, I came in contact with many fascinating students with whom I shared great companionships. Many thanks to Nick, Pat, Brandon, Sung, Joann, Dan, Leona, Angela, Dayu, Albert, Eric, and Will for creating a great atmosphere. In addition, I want to thank the students in Professor Lo’s optoelectronic group, including Victor, Jae, Eugene, Kai, Jessica, Randy, Sung-Hwan, Frank, Arthur, and Nicole for the productive collaboration.

I am also grateful to the advisors and coordinators in the bioengineering graduate student affairs office. Their tremendous effort in helping me to comply with the timeline and program requirements really made life easy for me. I would like to thank Irene Jacobo, Jeff Sanchez and Christine Miller for their work.

Last but not the least, I would like to acknowledge my family for their continuous support and encouragement.
VITA

Education:

• University of California, San Diego
  Major in Engineering Physics. 1994-1996
  Bachelor of Science in Bioengineering, June 1999

• University of California, San Diego
  Master in Science of Bioengineering, Dec 2004

• University of California, San Diego
  Doctor of Philosophy in Bioengineering, Mar 2007

Honors:

• Tau Beta Pi National Engineering Honor Society, 1999
• Golden Key National Honor Society, 1997

Research Experience:

• Dept of Bioengineering, UC San Diego (2001-Current)
  Graduate Research Assistant: Studying the role of Erythrocyte
  Tropomodulin (E-Tmod) in cardiac myofibrils. Research includes work on
  embryonic organ culture, microarray gene expression profiling of E-Tmod-
  deficient heart tissues, in vitro mechanical stimulation of cardiomyocytes
  and development of analytical software for microarray data. Ph.D. Co-
  advisors: Professors Shu Chien and Lanping Amy Sung.

• School of Medicine, UC San Diego (1998-2000)
  Undergraduate Research Assistant: Studying the protective effect of heat
  shock proteins in heart tissues against ischemia and reperfusion injuries.
  Mentors: Drs. Kurt Lin and Wolfgan Dillmann.

Teaching Experience:

• Bioengineering/Mathematics/MAE/ Depts, UC San Diego (2000-2006)
  Teaching Assistant:
  MAE 5 (Quantitative Computer Skills): Fall 2000
  Math 10A (Calculus): Winter 2001
  BE 160C/203 (Metabolic Engineering): Spring 2001
  BE 1 (Introduction to Bioengineering): Summer 2001
  BE 100 (Biostatistics): Spring 2001
  BE 230B (Molecular Biology): Spring 2003
  BE 250A (Biomechanics): Winter 2005
  BE 112B (Advanced Biomechanics): Spring 2005
  BE 208 (Molecular Bioengineering Laboratory): Fall 2005, Fall 2006

• International Business Machines, Inc. Toronto, CANADA (6/1997 &
  9/1999)
  Intern, training Instructor for customers, responsible for 3900 Mainframe
  System, 3490E, FC3000, FC0, F01 and F11 Tape Storage Subsystems
  Product Training.
Publications:

1. Kurt M. Lin, Brian Lin, **Ian Y. Lian**, Ruben Mestril, and Wolfgang H. Dillmann “Combined and individual mitochondrial HSP60 and HSP10 expression in cardiac myocytes protects mitochondrial function and prevents apoptotic cell deaths induced by simulated ischemia-reoxygenation”. *Circulation* 103: 1787-1792, 2001

2. Sepideh Heydarkhan-Hagvall, Shu Chien, Sven Nelder, Yi-Chen Li, Suli Yuan, Jianmin Lao, Jason H. Haga, **Ian Lian**, Phu Nguyen, Bo Risberg and Yi-Shuan Li “DNA Microarray Study on Gene Expression Profiles in Co-cultured Endothelial and Smooth Muscle cells in Response to 4- and 24-hour Shear Stress”. *Molecular and Cellular Biochemistry*, 281:1-15, 2006


5. **Ian Lian**, Andrea Bryan, Jeffrey Omens, Shu Chien and Lapng Amy Sung “Characterization of Cardiac Gene Expression and Phenotype of Erythrocyte Tropomodulin Knockout Mice”, in preparation

Selected Conference Abstracts:


**Ian Y. Lian** and Isaac Chu: “Integration of Novel Algorithms for Microarray Data Processing”. LabAutomation January 30-February 3, 2005, San Jose, CA

Yao, W., **Lian, I.,** Katz, B., and Sung, LA. Splice variants and alternative promoters of erythrocyte tropomodulin in erythroid cells under mechanical regulation. The 6th Annual UC Systemwide Bioengineering Symposium, June 25-27, 2005 Santa Cruz, CA

ABSTRACT OF THE DISSERTATION

The Roles of E-Tmod in Mechanotransduction and Cardiac Myofibrillogenesis

By

Ian Yu-Zen Lian

Doctor of Philosophy in Bioengineering

University of California, San Diego, 2007

Professor Shu Chien, Chair

Professor L. Amy Sung, Co-chair

Sarcomeres comprise the backbone of striated muscles, serving structural roles and providing contractile drive essential to cardiac functions. Among the components of the actomyosin complex that form the building blocks of sarcomeres, Erythrocyte Tropomodulin (E-Tmod) belongs to a special category of capping proteins known to regulate actin monomer polymerization.

While $E{-}Tmod^{-/-}$ is embryonically lethal with the phenotype of severe cardiac underdevelopment and impaired contractility, $E{-}Tmod^{+/-}$ mice survived through adulthood with a relatively mild phenotype of reduced right ventricular (RV) size. Through organ culture of the day post coitus (dpc) 9.5 $E{-}Tmod^{-/-}$ mouse hearts, we established that E-Tmod is required to sustain, but not to initiate, the heartbeats. Expression analysis also revealed that E-Tmod knockout caused the down-regulation of $\alpha$-Tropomyosin ($\alpha$-TM) and troponin, which are essential genes for $\text{Ca}^{2+}$ dependent contractile function. This finding suggested that E-Tmod influences cardiac contraction through modulating the TM-troponin complex.

Hypertrophy is common adaptive response of postnatal cardiomyocytes subjected to stress overload, which requires the synthesis of additional sarcomere
components to increase the cell size. Hypoxia stress overload remolds the underdeveloped RV of the $E$-$Tmod^{+/−}$ mice to the size of the wildtypes. Mechanical stretch and shear stress also elevated E-Tmod in both wildtype and siRNA E-Tmod knockdown myocytes. These results demonstrated that E-Tmod is up-regulated during hypertrophic process and that partial reduction of the E-Tmod (both $E$-$Tmod^{+/−}$ and siRNA-treated cardiomyocytes) does not affect the ability of the cardiomyocytes to undergo mechanically induced adaptive hypertrophy.

We have provided the direct expression evidence for the critical role of E-Tmod in the correlation between developmental evolution of ventricular diastolic function and changes in ventricular myoarchitecture. The present investigation also showed for the first time that alteration in E-Tmod directly affects the expressions of TM, actin, and troponins during myofibrillogenesis, as well as mechanical stretch induced hypertrophy. Therefore, our study suggests that E-Tmod is not simply a passive structural molecule, but plays important roles in the transcriptional regulation of multiple components of the sarcomeric thin filament.
Chapter I

Introduction

A. The Building Blocks of the Cardiac Contractile Machinery

Since Sir William Harvey (1578-1657) discovered the driving machinery of blood circulation in 1628 (An Anatomical Study of the Motion of the Heart and of the Blood in Animals), the heart has been recognized as the most physically energetic organ in the body with continuous life-long activity. An average human heart pumps over 7,000 liters of blood along 1,000,000 miles in the blood circulation daily, and contracts over 3 billion times over the course of the lifespan (Nicholas J. Severs 2000, BioEssays).

Myocardium is the muscular tissue of the heart that provides the driving force for cardiac contraction. Bounded by the epicardium (outer layer) and endocardium (inner lining), the myocardium is composed of cardiac muscle cells known as cardiomyocytes, which occupy around 80% of the total cell volume and account for more than 75% of the mass in mammalian hearts (Ding et al., 2003).

Myocardium contains myofibrils that are striated as in skeletal muscle, and a large portion of the internal volume of the cardiomyocytes is occupied by mitochondria, which synthesize ATP to supply energy for the constantly working cardiac muscle. Myofibrils and mitochondria occupy about 85% of the volume of cardiomyocytes, with the remainder occupied by the sarcolemma, T-tubules, sarcoplasmic reticulum, and specialized structures such as the intercalated disks. Cardiomyocytes make contact with adjacent cells via the gap junction, which makes connections through the plasma membranes of adjacent cells.
Cardiac contractility is driven by a protein machinery that consists of a nearly crystalline array of closely packed cytoskeletal proteins organized into a sarcomere. These sarcomeres are approximately 2 µm long and join in series by z-disks to form cardiac myofibrils. An adult human muscle fiber is typically 10-100 µm in diameter and can reach to cm in length (Craig and Padron, 2004).

The sarcomere structure exemplifies an effective and elegant design of the nature. The sliding members in each sarcomere unit consist of polarized thin actin filaments attached symmetrically to opposite ends of a bipolar thick myosin filament through myosin crossbridges (Fig. 1). During contraction, both filament types maintain relatively constant lengths (approximately 1.0 µm for thin filaments and 1.6 µm for thick filaments) during contraction. The sarcomere is highly conserved in its structure and geometry among diverse species (Craig and Padron, 2004). While it is still a mystery how thin and thick filaments in striated sarcomeres are maintained at such constant lengths across the species, capping proteins at the ends of the actin filament are implicated to play pivotal roles in the regulation and stabilization of such features, especially in the very dynamic, constantly contracting myofibrils of the heart.

The objective of our investigation is to elucidate the structural coordination between sarcomeric components, specifically those around the actin filament-capping protein E-Tmod, one of the least studied molecules in myofibrils to date. Investigations on how members of the myofibrils interact and are regulated during development and under mechanical stimulations, the interrelationships between functions and expressions can elucidated which have broad biological significance.
B. Components of the Cardiac Sarcomere

Sarcomeres are multi-protein complexes composed of three different filament systems. The thick filament system is composed of the myosin protein. The thin filaments are assembled by actin monomers. The elastic filament system is composed of the giant protein titin (also known as connectin). Myosin filaments extend throughout the A-band and are thought to overlap in the M-band. Actin filaments are the major component of the I-band and extend into the A-band. The giant protein titin (connectin) extends from the Z-line of the sarcomere, where it binds to the thin filament system and to the M-band, where it is thought to interact with the thick filaments. Titin (and its splice isoforms) is the largest single protein found in nature. It provides binding sites for numerous proteins and is thought to
play an important role as sarcomeric ruler and as a blueprint for the assembly of the sarcomere.

Several proteins important for the stability of the sarcomeric structure are found in the Z-line as well as in the M-band of the sarcomere. Actin filaments and Titin molecules are cross-linked in the Z-disc via the Z-line protein alpha-Actinin. The M-band proteins myomesin and M-protein crosslink the thick filament system (myosins) and the M-band part of titin (the elastic filaments). The interaction between actin and myosin filaments in the A-band of the sarcomere is responsible for the muscle contraction (sliding filament model).

C. The Cardiac Myofibril Z-Disk

The Z-line is a multifunctional macromolecular complex that anchors sarcomeric actin filaments, mediates interactions with intermediate filaments and costameres, and recruits signaling molecules (Henderson et al. 2003). As an important part of the myofibril structure that gives the striated feature of skeletal and cardiac muscle, the cardiac muscle Z-line has the primary function of crosslinking anti-parallel actin filaments and transmit force from the adjacent sarcomeres. Z-lines distribute contractile forces through the sarcomeric matrix and act to maintain myofibril integrity (Wang et al., 1993; Li et al., 1997). In addition to the structural roles, the Z-line has also been shown to serve as a signaling node that recruits protein phosphatase and calcineurin (Frey et al., 2000).

D. Overview of the sarcomeric defects and cardiac phenotypes

D.1. Phenotypes of Sarcomere Mutations

Sarcomere abnormality is the major cause of cardiomyopathies (Vang, et al., 2005). Specifically, 55-70% of the hypertrophic cardiomyopathy (HCM) and 30% of the dilated cardiomyopathy (DCM) are identified with sarcomere mutations
The correlation between hereditary cardiac conditions and specific defects in the molecular components has been the focus of the genetic research for decades (Table 1). It has been reported that mutations that affect sarcomere contraction lead to HCM, whereas mutations that affect force transmission from the sarcomere to the surroundings (including neighbouring sarcomeres, extracellular matrix and cytoskeleton) often lead to DCM (Chien 2003; Mogensen et al., 2004).

### Table 1: Cardiac phenotype of sarcomeric mutations

<table>
<thead>
<tr>
<th>Genetic Defect</th>
<th>HCM</th>
<th>DCM</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiac actin</td>
<td>Missense</td>
<td>Missense</td>
<td>Olson, et al. 1998; Mogensen et al. 1999</td>
</tr>
<tr>
<td>Myosin heavy chain</td>
<td>Missense</td>
<td>Missense</td>
<td>Tanigawa et al. 1990; Bonne et al. 1998; Kamisago 2000</td>
</tr>
<tr>
<td>Myosin essential light chain</td>
<td>Missense</td>
<td></td>
<td>Poetter et al 1996</td>
</tr>
<tr>
<td>Myosin regulatory light chain</td>
<td>Missense</td>
<td></td>
<td>Poetter et al 1996</td>
</tr>
<tr>
<td>α-Tropomyosin</td>
<td>Missense</td>
<td>Missense</td>
<td>Bonne et al. 1998 Thierfelder et al. 1994 Olsen et al. 2001</td>
</tr>
<tr>
<td>Titin</td>
<td>Missense</td>
<td>Deletion, missense</td>
<td>Gerull et al. 2002 Itoh-Satoh et al. 2002</td>
</tr>
<tr>
<td>MLP</td>
<td>Missense</td>
<td></td>
<td>Knoll et al. 2002</td>
</tr>
<tr>
<td>Desmin</td>
<td>Missense</td>
<td></td>
<td>Goldfarb et al. 1998</td>
</tr>
<tr>
<td>Troponin I</td>
<td>Missense</td>
<td></td>
<td>Kimura et al. 1997</td>
</tr>
<tr>
<td>Lamin A/C</td>
<td>Missense</td>
<td></td>
<td>Fatkin et al. 1999</td>
</tr>
</tbody>
</table>

(adapted from Chien K, 2003)
D.2. Phenotypes of Sarcomere Deletions

While hereditary mutations are valuable cases to study the disease phenotypes of the affected sarcomeric genes, genetically engineered knockout (KO) mice can provide a definitive null model for *in vivo* studies. Among the genes encoding components of the cardiac myofibrils, all KO mouse models generated to date exhibit moderate to severe abnormalities in cardiac phenotypes (Table 2). Similar to the hereditary mutation phenotypes, KO of sarcomeric molecules results in various forms of cardiomyopathies in mice. In addition, many of those have underdeveloped chambers, paricardiac edema and die prematurely. Notably, mice with KO in actomyosin component of the sarcomere do not survive beyond the embryonic stage, but mice with absence of Z-disk proteins can live to adulthood. Since components of actin, myosin and titin filaments make up the building blocks of sarcomeres, while z-disk proteins provide an important link between structural proteins to mechanotransduction and cardiomyocyte signaling (Knoll *et al.*, 2002). Defects in sarcomeric components and regulation therefore lead to significant clinical challenges common to the cardiomyopathies.
Table 2: Cardiac phenotypes of sarcomeric knockouts

<table>
<thead>
<tr>
<th>Target Gene</th>
<th>Embryonic Lethal</th>
<th>Cardiac Phenotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>α cardiac actin</td>
<td>Yes</td>
<td>DCM, myofibril disarray. 40% survival postnatally but not beyond 2 weeks.</td>
<td>Kumar et al. 1997</td>
</tr>
<tr>
<td>Tropomodulin</td>
<td>YES (E10)</td>
<td>Underdeveloped heart w. pericardiac edema</td>
<td>Chu et al. 2003 Fritz-Six et al. 2003</td>
</tr>
<tr>
<td>Nebulin</td>
<td>NO</td>
<td>Severe muscle weakness w. reduced A band</td>
<td>Bang et al. 2006</td>
</tr>
<tr>
<td>Nebulette</td>
<td>NO</td>
<td>Loss of TM, reduced beating frequencies</td>
<td>Moncman and Wang, 2002</td>
</tr>
<tr>
<td>Desmin</td>
<td>NO</td>
<td>DCM, Z-line disarray</td>
<td>Milner et al. 1999</td>
</tr>
<tr>
<td>α- Tropomyosin</td>
<td>Yes (E8-E11.5)</td>
<td>HCM, myocyte disarray and asymmetric ventricular hypertrophy, thrombi formation in left atrium</td>
<td>Rethinasamy et al. 1997</td>
</tr>
<tr>
<td>Myosin MHC</td>
<td>Yes (E11-E12)</td>
<td>FMC (BMHC KO), atrial septation</td>
<td>Keith-Jones et al. 1996</td>
</tr>
<tr>
<td>Myosin MLC2a</td>
<td>Yes (E10.5- E11.5)</td>
<td>severe chest edema, growth arrest at approximately somite pair (sp) 24-28 with growth retardation begins at E9.5</td>
<td>Huang et al. 2003</td>
</tr>
<tr>
<td>Myosin MLC2v</td>
<td>Yes (E12.5)</td>
<td>massive cardiac enlargement, wall thinning, chamber dilation, and pleural effusions, hepatic congestion and an engorged vena cava, MLC2a compensation</td>
<td>Hunag et al. 2003</td>
</tr>
<tr>
<td>Titin</td>
<td>Yes (E10-12)</td>
<td>M-Line connection to Z is lost. Loss of sarcomeric striation;</td>
<td>Gotthardt et al. 2003 Weinert et al. 2006</td>
</tr>
</tbody>
</table>

E. Introduction to the Tropomodulin gene family

Tmod is a class of proteins that express and provide essential functions across a variety of mammalian tissues. Among the four known Tmod homologues,
erythrocyte tropomodulin (E-Tmod) was the first one discovered (Fowler, 1987). Originally cloned from human reticulocyte and fetal liver cDNA libraries (Sung et al., 1992), E-Tmod is classified as a structurally critical tropomyosin (TM)-binding, and actin-capping protein that is expressed in abundance in murine cardiac and slow-twitching skeletal muscles (Conley, 2001).

E-Tmod is one of the few capping proteins localized in the pointed-end actin filaments, in contrast to the many known to cap the barbed end of the filaments, such as adducin, CapZ, and gesolin (Fowler, 1997). Capable of capping the slow-growing end of the actin filament and binding to the N-terminal of the tropomyosin molecule (Sung et al., 1994), E-Tmod plays an important role in regulating the actin polymerization/depolymerization process. My working hypothesis is that this regulatory mechanism is carried out via the capping function of the E-Tmod; this hypothesis is supported by the decrease in the affinity between actin monomers and the pointed ends of the thin filaments following E-Tmod over-expression, with monomer dissociation and filament shortening (Littlefield et al., 2001).

After the initial discovery of E-Tmod, three additional Tropomodulin homologues have been identified within the past decade. The four members of the Tmod family share high degree of peptide sequence homology and are highly conserved among vertebrates, and all Tmod variants possess the common ability to bind and cap the barbed (slow growing) end of the actin filament among various tissue types.

Among four members of the Tmod family, Tropomodulin 2, known as neural-specific N-Tmod and as neural-specific Tmod-2, is expressed exclusively in the brain, spinal cord and peripheral nerves. N-Tmod−/− mice are viable and fertile, and exhibit no anatomical abnormalities, but behavioral analysis showed hyperactivity,
reduced sensorimotor gating and impaired learning and memory, suggesting that N-Tmod plays a role in behavior, learning, memory, and synaptic plasticity (Cox et al., 2003). Studies on the N-Tmod knockout mice reported an increase transcriptional expression of E-Tmod levels in both heterozygous (2 fold) and null (8 fold) knockouts in the brain tissue, suggests a compensatory response of E-Tmod to N-Tmod deficiency (Cox et al., 2003).

U-Tmod, the ubiquitous Tmod, is expressed in many tissue types and especially abundant in endothelial cells (Verrills et al., 2003). Recently, a novel hypothesis has been proposed, in which pointed-end capping of actin filaments by U-Tmod can negatively regulate endothelial cell motility (Fischer et al., 2003). This theory is supported by experimental evidence that transient over-expression of GFP-U-Tmod leads to decreased cell motility, and that reduction of U-Tmod by RNA interference results in faster average cell migration, accompanied by increases in free pointed and barbed ends in lamellipodial actin filaments. These findings suggest that the tropomodulin capping function is not limited to pointed ends of actin filaments in stable, nondynamic cytoskeleton structures.

Sk-Tmod (skeletal tropomodulin, Tmod-4) is predominantly expressed in fast skeletal muscle fibers, lens, and erythrocytes, but not in cardiomyocytes of mouse and human (Almenar-Queralt et al., 1999). While co-expressed in skeletal muscles with E-Tmod, Sk-Tmod is localized on regions of the cytoskeletal structure in myofibrils, rather than the costameric region where E-Tmod is found. Sk-Tmod also uniquely binds to the lens-specific intermediate filament protein, filensin (Fischer et al., 2003). Being expressed with low affinity and stoichiometry, Sk-Tmod does not inhibit the actin-capping activity in vitro, suggest that Sk-Tmod serves additional non-capping related functions which are still unknown.
F. Significance of Erythrocyte Tropomodulin in the Heart

In mammals and many other animals, the heart is the first organ to form during the embryonic stage and engages in contraction prior to its full development. Presence of E-Tmod can be detected in the embryonic mouse heart prior to and throughout the myofibril maturation (Fig 2). Absence of E-Tmod result in embryonic lethality around E10 (Chu et al., 2003).

E-Tmod null (-/-) knockout mouse exhibits severe cardiac phenotypes, including underdeveloped ventricles, pericardiac edema, and impaired contractility (Chu et al., 2003). Structurally, myofibrils are disorganized without striation, and the H-zone gaps in F-actin staining are not observed (Fritz-Six et al., 2003). Since the maturations of myofibrils and mechanical properties of sarcomeres are critical for heart morphogenesis during embryonic development, E-Tmod is implicated to play a pivotal role in cardiac function and development.

![Figure 2. Localization of E-Tmod in cardiac myofibrils during cardiogenesis. F-Actin (Red), E-Tmod (Green). Bars = 10 µm](image)

G. Hypothesis and objectives

We hypothesize that the role of E-Tmod in actin capping has significant influence on cardiomyocyte structure and function. Specifically, it may govern the organization of actin filaments, which affects the load regulation in
mechanotransduction through the cytoskeletal organization. The goals of the current study are to characterize the cardiac phenotype of the E-Tmod knockout mice, and to profile the differentially expressed genes in the heart, with the aims of elucidating the potential links between the altered genes and the affected phenotype. In addition, we will investigate the impact of mechanical stimuli and myofibrillogenesis on the regulation of actin and tropomyosin in relation to the E-Tmod expression.

The combined approach of high-throughput gene expression profiling, custom-developed analytical algorithms, molecular cloning, and protein assays, in conjunction with the knockout mouse system, will be fully explored in this study. Specifically, we will (1) establish the temporal and spatial expression pattern of the E-Tmod during cardiac development. (2) clone and identify cardiac-specific E-Tmod splice variants in mouse and rat models. (3) characterize the cardiac phenotype in E-Tmod KO mouse heart, and to profile the global-scale gene expression in the KO hearts with DNA microarrays and complementary technologies. (4) explore and develop bioinformatics algorithms to cluster and map high-throughput expression data to functional biological pathways. (5) elucidate the effects of different types of mechanical stimuli on the phenotype of cardiomyocytes and expression pattern of the corresponding E-Tmod variants. The basis of such knowledge will provide novel information regarding the role of E-Tmod in modulating sarcomere and cytoskeletal structure and function under mechanical stimuli and during myofibrillogenesis.
Chapter II

Gene Expression Profiling of the E-Tmod deficient Mice

Abstract

Genetic knockout mouse is often used as an in vivo platform to elucidate the function of the disrupted gene. Normal murine cardiac tissues express a high level of erythrocyte tropomodulin (E-Tmod), a capping protein for actin filaments at the pointed end. While homozygocity (E-Tmod^{−−}) was previously found to be lethal around embryonic stage E10, heterozygous (E-Tmod^{−+}) mice can survive, but develop smaller right ventricles. Surveying of gene expressions in E-Tmod^{−+} cardiac tissues with DNA microarray revealed a number of gene clusters being consistently differentially expressed; notably the down-regulation of multiple members in the cyclin, myogenic and several actomyosin gene families. Furthermore, the regulatory trends of these genes are in excellent agreement between the E-Tmod^{+-} and E-Tmod^{−−} tissues, providing a mechanistic explanation for the E-Tmod deficient cardiac phenotypes based on the genetic alternations.

Introduction

Defects in cytoskeletal and sarcomeric genes have been associated with various forms of cardiomyopathies (Chien, 2003). Genetic knockout studies showed that while disruptions in actomyosin molecules are embryonically lethal, deficiencies in Z-line components are often survivable (Table 2, Chapter 1). These findings suggest that elements of contractile filaments are more critical in function and
development relative to the non-contractile sarcomere components whose manifestation is to transmit force.

Among various cytoskeletal and sarcomeric molecules, erythrocyte tropomodulin (E-Tmod) is one of the less-studied components of the actomyosin complex. Serving as a “capping” protein at the pointed (slow-growing) end of the actin thin filament, E-Tmod has been shown to interact with actin (Fowler et al., 2003), tropomyosin (Sung et al., 1994), and possibly nebulin (Witt et al., 2006) with varying degrees of affinities.

Altered E-Tmod expressions have been shown to cause significant abnormalities in various animal models. In vitro overexpression of E-Tmod in both rat and chick cardiomyocytes contributed to the shortening of thin filaments and disorganization of sarcomeres, resulting in myofibril degeneration (Littlefield et al., 2001). Likewise, in vivo overexpression of E-Tmod in the mouse heart led to disrupted sarcomere organization with shortened thin filaments, resulting in myofibril degeneration and dilated cardiomyopathy (Sussman et al., 1998). Absence of E-Tmod vitally affected the contractility of the ES-derived cardiomyocytes (Ono et al., 2005). A genetic knockout of E-Tmod alleles in mice resulted in lethality at approximately embryonic day 10 (E10). Phenotypically, E-Tmod−/− embryos showed distorted heart growth, disrupted vessel development, and restricted erythrogenesis in the yolk sac (Chu et al., 2003). These phenomena demonstrated a critical importance of E-Tmod in the development of cardiovascular and possibly other systems.

It is interesting that, while mice with knockout of the actin-capping protein at the slow-growing end (E-Tmod) do not survive (Mills et al., 2001), mice with knockouts of the more dynamic fast-growing (Cap-Z) end can survive (Pyle et al.,
These findings suggest actin-capping activities at two ends of the thin filament do not equal. More importantly, E-Tmod cannot be replaced nor compensated by other structural molecules as oppose to the barbed-edn capping protein capZ and β-adduin.

Genetic knockout model often provides an excellent platform for the understanding of a targeted disruption gene. A recent study on β-adducin (capping protein at the fast growing end of the actin thin filament) revealed a compensatory upregulation of CapZ and E-Tmod, accompanied by the down-regulation of α-tropomyosin (α-TM) and actin (Porro et al., 2004). These findings suggest a strong regulatory association among some genes encoding for the actin-filament complex, specifically within members of the capping protein family. By investigating the differential coexpression of E-Tmod deficient mice, gene clusters with altered expressions due to E-Tmod deficiency can be identified. Such data can provide an understanding of the transcriptional basis for the phenotypes observed in the KO, and this would have significant implications in cardiac physiology relevant to sarcomeric defects related to E-Tmod abnormalities.

**Material and Methods**

**A. Sectioning and staining of whole heart tissues**

A Leica CM3050 Cryostat was utilized to cut 20-µm thick frozen sections every 400 µm throughout each heart. These short-axis sections were taken perpendicular to the long axis of the heart defined by the aorta-apex line. Sections were stained using a solution of picric acid (100 ml) and sirius red (0.1 g), a collagen specific stain. The sections were examined with a Nikon OptiPhot-2 microscope with a low-power objective, digitized with a Nikon CoolPix 990 digital
camera, and analyzed using the NIH imaging software (http://rsb.info.nih.gov/nih-image/about.html). After image enhancement, the software was used to calculate the right ventricular chamber area and the right ventricular free-wall area from each slice. Simpson’s trapezoid method of discs was employed for calculating the right ventricular chamber volume in each heart.

**B. Strategy for high-throughput expression profiling of E-Tmod<sup>+/−</sup> cardiac tissue**

Based on a mouse oligoneucletide array designed and constructed according to the description in the Material and Method section E, the strategy to profile the large-scale gene expression in the E-Tmod<sup>+/−</sup> hearts was devised as shown in figure 3. The design of the project included the n = 6 array experiments comparing samples extracted from 6 pairs of age and weight matched mice (each pair comprised of a wildtype control and an E-Tmod<sup>+/−</sup> mice), complemented by the independent validation with the low-throughput but highly sensitive qPCR assays.

Figure 3. Strategy of large-scale gene expression profiling scheme utilizing the combination of microarray and qPCR platforms.
C. Collection of mouse hearts

Twelve (six $E\text{-}Tmod^{+/−}$ and six $E\text{-}Tmod^{+/+}$) weight-matched male mice at 40 weeks of age were used for this study. The mice were sacrificed by cervical dislocation in compliance with Animal Care Program (ACP) approved protocol (Animal care during all experiments was approved by the University of California, San Diego, Animal Subjects Committee in accordance with National Institutes of Health guidelines). After recording of body weight, the heart was exercised, sliced open, rinsed in PBS, weighted, and immediately subjected to RNA extraction.

D. Heart RNA and cDNA preparations

The cardiac tissues were homogenized immediately after harvesting and the RNA pellet was extracted in Trizol™ (GIBCO, Paisley, UK) solution using a protocol supplied by the manufacturer. The RNA was then resuspended in nuclease-free DEPC-treated water. The quality of RNA was examined by agarose gel electrophoresis using the same amount of cardiac RNA from each animal, and by confirming the OD 260/280 ratio to be at least 1.8. Thirty µg of total RNA (3 µg/µl) from each animal was reverse transcribed at 42°C for 1.5 hr by using the CyScribe First-Strand cDNA Labeling Kit (Amersham Biosciences, Buckinghamshire, UK). cDNA synthesis was primed with the oligo-dT primer, which anneals to the poly A-tails of the transcripts (only accounting for 1.5-2% of the total RNA) to selectively prime cDNA synthesis from the source RNA. As this method only created one transcription start site per transcript, unincorporated nucleotides were removed by using a MicroSpin G-50 Micro column (Amersham Biosciences, Piscataway, NJ).
E. Mouse oligonucleotide microarray

A sequence-verified mouse 70-mer oligonucleotide library (Mouse Oligo Set Version 1.1) containing 6,912 unique genes was obtained from Operon Technologies (Alameda, CA). The array-ready, melting-temperature optimized DNA solution (600 pmol) was transferred to 384-well plates for array spotting. The library design is based on the Ensembl (http://www.ensembl.org/) Mouse 14.30 Database and Mouse Genome Sequencing Project. The oligo set was accompanied by complete annotation, including Gene Ontology (GO) and oligo chromosome coordinates. The spotting was carried out in SpotArray 24 (Perkin Elmer Life Sciences Inc., Boston, MA), using a 24-pin format, onto CMT-GAPS II-coated slide (Corning Co., Corning, NY) under 50% relative humidity. The printed slide was baked for 2 hrs at 80°C and stored in desiccators at room temperature. Spotting quality was examined by hybridization to Cy3-labeled vector sequence. Point-to-point fluctuation between slides was under 5% judging by the fluorescence intensity.

F. Microarray pre-hybridization and hybridization

The microarrays were prehybridized at 42°C for 45 min in 50 ml of a solution containing 1% BSA, and 0.2X saline-sodium citrate (SSC, which contains 1% 150 mM NaCl, 1% polyvinylpyrrolidone, 5 mM sodium citrate, and 1% glycine at pH 7.0) (Soares et al., 1994). cDNA target preps were concentrated to 5 µl or less by vacuum-drying (Savant Speed Vac Model 110, Farmingdale, NY). After adding DIG Easy Hyb solution (Roche Applied Science, Indianapolis, IN) to a total volume of 35 µl, the hybridization solution was placed evenly onto the custom-spotted oligonucleotide chip, enclosed within a sealed hybridization chamber at 42°C for 12-hour target-probe hybridization.
**G. Scanning and image quantification**

The hybridized arrays were washed with 2X SSC twice, and 0.2X SSC once. After drying by air, the slides were scanned with a ScanArray 4000 Microarray Scanner (GSI Lumonics/Packard Instruments, Union City, CA) with adjustable laser and photo multiplier tube (PMT) settings. The measured intensities of the arrayed spots are transformed into quantified expression values. The image acquisition software ScanArray and image quantification tool, including QuantArray as well as GenePix, were utilized to complete the spots quantification tasks.

**H. Designation of flag (low quality) spots**

Individual spots on the scanned array were screened for physical defects, including morphology, size, alignment and signal-to-noise ratio (S/N). The cutoff criteria for flagging were set at a circularity score ≤ 50%, spot size deviation ≥ 50%, alignment gap ≥ 100% diameter, and S/N ≥ 3.

**I. Data mining and analysis**

In an effort to streamline and improve the microarray data analysis, an integrated software package with interactive interface platform was developed and applied to the experimental data as shown in the scheme (Fig 4). In brief, background signals of the numerically quantified raw data were smoothened with a weighted moving-average algorithm before being subtracted from the hybridization signals. The interim data were then subjected to log-based global normalization to remove dye bias, followed by a refined local normalization procedure that minimizes intensity related bias. The gene expression ratio was automatically tabulated and stored from each experiment, and cross-array normalization was carried out to match the detection range of each array, and the
final gene expression profile was outputted from replicated experiments.

Figure 4. Flowchart of the application of improved-algorithms to process microarray data. Proprietary data treatment scheme were developed and validated (sections H-J) to show (A) improved S/N by minimizing the noise level, (2) reduction of outliers, and (3) removal of dye- and array-related bias.

**J. Defining differentially expressed (DE) genes**

Normalized data were screened and sorted by expression patterns, functional categories, and then filtered by t-test (P value < 0.05) and coefficient of variations (< 50%) to guard against random or statistically inconsistent results. The criteria for “differentially expressed” genes were defined by “deviation distribution tier” instead of conventional numerical cutoff method (e.g. two-fold change) to avoid subjective and arbitrary standard. The genes were functionally categorized and clustered by GeneSet Enrichment Analysis (GSEA) method according to the definitions of the Gene Ontology (GO) Consortium (http://www.geneontology.org/). Genes that did not meet the criteria above were omitted from further analysis, but are included in the appendix online (http://be-web.ucsd.edu/faculty/area/chien_lab/personnel.html#lian).
K. Quantitative real-time PCR reactions

The RNA sample (100 ng) and 50 pmol of the corresponding forward and reverse primers (Qiagen, Santa Clarita, CA) were mixed together with 25 µl of SYBR Green Super Mix, according to the manufacture’s protocol (BioRad, Hercules, CA), to yield a total volume of 50 µl. PCR primers (Table 3) were designed and optimized using the MacVector program (Accelrys, San Diego, CA), and searched against the public database to confirm unique amplification products (http://www.ncbi.nlm.nih.gov). Primers ranging from 20 to 25 nucleotides were chosen to validate the gene expressions found by microarray tests to generate PCR products of 109 to 161 base pairs. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH), which is a catalytic enzyme involved in glycolysis, was used as the control for the qPCR assay. Reactions in 96-well format were performed in the single color real-time PCR detection system (MyiQ, BioRad, Hercules, CA). The cycling parameters were 95°C for 15 min, followed by 40 cycles of PCR (15 sec at 94°C and 1 min at 60°C). The threshold cycle value (C<sub>T</sub>) represents the cycle at which a statistically significant increase in the normalized reporter signal (R<sub>n</sub>) above a chosen threshold can first be detected, according to the manufacturer’s manual. Threshold was defined as the average standard deviation of R<sub>n</sub> for the early cycles, multiplied by a normalization factor collected from all sample wells. The relative RNA expression of each sample was determined from the experimental curve normalized to the control (GAPDH) curve and then to its corresponding wildtype counterpart.
Table 3. Primers for qPCR and the size of expected products

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-Actin</td>
<td>Forward: 5′- CCAGTTGGTATAATGGCCATGTC -3′</td>
</tr>
<tr>
<td>(142 nt)</td>
<td>Reverse: 5′- TCTCTAGTGAGGATGAGAGCAG -3′</td>
</tr>
<tr>
<td>BMP 5</td>
<td>Forward: 5′- CTTTTGCACAGGAAACTACAG -3′</td>
</tr>
<tr>
<td>(136 nt)</td>
<td>Reverse: 5′- ATAGATGATGCTGATGAGAGCAG -3′</td>
</tr>
<tr>
<td>CDK 5</td>
<td>Forward: 5′- TCAACAAATAATATTCGTCGTC -3′</td>
</tr>
<tr>
<td>(121 nt)</td>
<td>Reverse: 5′- AAATTTCGTCGTCGTCGTCGTCG -3′</td>
</tr>
<tr>
<td>CDK 6</td>
<td>Forward: 5′- TCTATTGAGGGATTTCGTCGTC -3′</td>
</tr>
<tr>
<td>(103 nt)</td>
<td>Reverse: 5′- AAATTTCGTCGTCGTCGTCGTCG -3′</td>
</tr>
<tr>
<td>cyclin D1</td>
<td>Forward: 5′- GCAATGGAACCTGCTGCTGCTG -3′</td>
</tr>
<tr>
<td>(132 nt)</td>
<td>Reverse: 5′- GAGATGTTGCGCAAGAGGTCGTC -3′</td>
</tr>
<tr>
<td>cyclin D2</td>
<td>Forward: 5′- GCTGTCGATTTGCCGCAACGTCG -3′</td>
</tr>
<tr>
<td>(161 nt)</td>
<td>Reverse: 5′- CTTGCCGAGGATGTCAGTCGTCG -3′</td>
</tr>
<tr>
<td>cyclin E1</td>
<td>Forward: 5′- CTGTGAAAGCCGATGAGCAGGTC -3′</td>
</tr>
<tr>
<td>(93 nt)</td>
<td>Reverse: 5′- CTGGGGATGATGCTGCTGCTGCTG -3′</td>
</tr>
<tr>
<td>cyclin E2</td>
<td>Forward: 5′- CCTTGGTCGTCGTCGTCGTCGTCG -3′</td>
</tr>
<tr>
<td>(103 nt)</td>
<td>Reverse: 5′- CTTGCCGAGGATAAAATGCTGTCG -3′</td>
</tr>
<tr>
<td>cyclin F</td>
<td>Forward: 5′- GGGGATGCTGCTGCTGCTGCTGCTG -3′</td>
</tr>
<tr>
<td>(171 nt)</td>
<td>Reverse: 5′- CTTGCCGAGGATAAAATGCTGCTGCTG -3′</td>
</tr>
<tr>
<td>cyclin G1</td>
<td>Forward: 5′- GAGATGTTGCGCAAGAGGTCGTCG -3′</td>
</tr>
<tr>
<td>(83 nt)</td>
<td>Reverse: 5′- CTTGCCGAGGATAAAATGCTGCTGCTG -3′</td>
</tr>
<tr>
<td>cyclin G2</td>
<td>Forward: 5′- TCTTGGTCGTCGTCGTCGTCGTCG -3′</td>
</tr>
<tr>
<td>(112 nt)</td>
<td>Reverse: 5′- CTTGCCGAGGATAAAATGCTGCTGCTG -3′</td>
</tr>
<tr>
<td>E-Tmod</td>
<td>Forward: 5′- CACCATGTCCTAGGACGAGGTCG -3′</td>
</tr>
<tr>
<td>(152 nt)</td>
<td>Reverse: 5′- CACATTGTCCTAGGACGAGGTCG -3′</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Forward: 5′- GCCTGGTATGCTGCTGCTGCTGCTG -3′</td>
</tr>
<tr>
<td>(137 nt)</td>
<td>Reverse: 5′- GCCTGGTATGCTGCTGCTGCTGCTG -3′</td>
</tr>
<tr>
<td>TroponinI</td>
<td>Forward: 5′- TGAGTTGGAAGCGAAAATGTCG -3′</td>
</tr>
<tr>
<td>(137 nt)</td>
<td>Reverse: 5′- CAATTGTCCTGCTGCTGCTGCTGCTG -3′</td>
</tr>
<tr>
<td>TPM 2</td>
<td>Forward: 5′- CAATGGAACCGCTGCTGCTGCTG -3′</td>
</tr>
<tr>
<td>(121 nt)</td>
<td>Reverse: 5′- CAATGGAACCGCTGCTGCTGCTGCTG -3′</td>
</tr>
</tbody>
</table>

Results

A. Cardiac phenotype of E-Tmod<sup>−/−</sup> and E-Tmod<sup>+/+</sup> mice

E-Tmod<sup>−/−</sup> were embryonically lethal and displayed characteristically smaller hearts embedded in the pericardial edema by the gestational age of E9 – E10 (Fig. 5A). Ventricles of E-Tmod<sup>−/−</sup> embryos, especially the right ventricles (RV), were clearly underdeveloped (Fig. 5B). Cardiomyocytes of the E-Tmod<sup>−/−</sup> samples were smaller in size and exhibited fuzzy, disarray cellular boarders compared with the wildtypes (Fig. 5C) collected from the same developmental stage. Cardiac sections of the adult E-Tmod<sup>+/+</sup> mice revealed the reduction of both right RV cavity (by 66%) and wall volume (by 26%) relative to the wildtype controls (Fig. 6).
Figure 5. Morphology of the $E$-$Tmod^{-/-}$ embryo proper (A), whole heart (B) and cardiomyocyte staining (C) at gestational age E9. Images (A) and (B), whole mount in situ hybridization on embryos using antisense cRNA probes was presented in $E$-$Tmod^{+/+}$ (top) and $-/-$ (bottom), are adapted from Chu et al., 2003. All images are front views except for eHAND, which is a left front view. h, heart; v, ventricle; lv, left ventricle; rv, right ventricle; ot, outflow tract; pe, pericardial effusion.
Figure 6. RV cavity and free wall volumes of $E-Tmod^{+/−}$ and wildtype mice. The volumes were calculated by compiling the areas of the 20 µm-thick cardiac sections prepared according to the procedures described in material and method.

**B. E-Tmod expressions in $E-Tmod^{−/−}$ and $E-Tmod^{+/−}$ embryonic hearts**

The E-Tmod expressions in the embryonic hearts of the heterozygous and homozygous knockouts were determined using quantitative PCR assays. As expected, the $E-Tmod^{−/−}$ samples showed no detectable level of E-Tmod using exon 1 as the forward primer. $E-Tmod^{+/−}$ expressed 51% of the transcript (primed by E1 and E9 as forward and reverse primer respectively) relative to the wildtype (Fig. 7), indicating that the heterozygous knockout contained approximately one-half of the basal E-Tmod level.
Figure 7. (A) Expression of E-Tmod in $E-Tmod^{-/-}$, $E-Tmod^{+/+}$ and $E-Tmod^{+/-}$ embryo hearts. qPCR reactions were primed by the exon 1 and exon 9 of the E-Tmod transcript yielding the 1,077 bp products confirmed by the agarose gel (B). The cDNA were extracted from cardiac tissues of the embryonic samples collected at E9.5 of the gestational age. Expression levels in (A) are normalized to the internal control GAPDH.

C. Expressions of Tmod homologs in $E-Tmod^{+/-}$ and $E-Tmod^{-/-}$ hearts

The gene expression assay of four Tmods showed $U-Tmod$ to be the most abundantly expressed Tmod homologue in the embryo propers, followed by $N-Tmod$, $E-Tmod$ and $Sk-Tmod$ (Fig. 8). Interestingly, the expression of $U-Tmod$ is twice as much in the $E-Tmod^{+/-}$ and $E-Tmod^{-/-}$ E9.5 embryo propers as that in the $Tmod^{+/-}$, suggesting a novel compensatory response of $U-Tmod$ to $E-Tmod$ knockout in these embryos.
Figure 8. Tmod homolog expression profile in \(E-Tmod^{-/-}\), \(E-Tmod^{+/+}\) and \(E-Tmod^{++}\) E10 embryo proper. The % value indicates the proportion of individual Tmod homolog contribution, which summed up to 100% in each genotype group.

D. Number of differentially expressed genes in \(E-Tmod^{+/-}\) cardiac tissues

Microarray data collected from 6 pairs of cardiac samples (each pair included a wildtype control and an \(E-Tmod^{+/-}\) tissue) identified the differentially expressed genes from the heterozygous E-Tmod KO mice. Within 6,912 genes being profiled, a total of 1,395 members are found to be differentially expressed with statistical significance. Among those genes, 410 were up-regulated and 985 were down-regulated (Fig. 9). Specifically, 60 out of the 410 up-regulated and 669 out of the 985 down-regulated gene have satisfied triple statistical criteria described in material and method section J.
Figure 9. Venn diagrams summarizing the number of differentially expressed genes with respect to their statistical properties in the E-Tmod cardiac tissue compared to the wildtype. The numbers that overlap among 3 statistical criteria describe are shown inside the circles.

E. Clusters of differentially expressed genes in E-Tmod$^{+/−}$ cardiac tissues

The list of down-(table 4A) and up-(table 4B) regulated genes in cardiac tissues of E-Tmod$^{+/−}$ mice were generated using the criteria described in Material and Method section J. Members of actin ($α$-cardiac actin), tropomyosin ($β$-TM), and actin-binding (cappa3) genes are found to be down-regulated in E-Tmod$^{+/−}$ cardiac tissue. Many of myosin genes are found to differentially expressed, including the up-regulation of myosin heavy peptide 7 (Myh7), myosin 1F and myosin 10. In addition, a group of genes from the cyclin and cyclin-dependent kinase (CDK) families are down-regulated in E-Tmod$^{+/−}$ hearts. A few myogenic and differentiation factors, including genes encoding bone morphogenic proteins (BMP5, BMP7 and BMP10) and myogenic differentiation 1 (MyoD1), were also among the differentially expressed genes. Cluster analysis (Fig. 10) based on expression shows that “development and organogenesis” and “cell cycle” are among the top three categories of genes down-regulated in the E-Tmod cardiac tissue, while “DNA metabolism and regulation” being the top up-regulated gene group.
Figure 10. Distribution of the down-regulated (A) and up-regulated (B) DE genes passing three criteria (P < 0.05, CV < 50% and agreement > 66% among replicates).

Table 4A. Selected down-regulated genes from the E-Tmod−/+ cardiac tissue

<table>
<thead>
<tr>
<th>Ave Ratio</th>
<th>SB</th>
<th>CV</th>
<th>P-Value</th>
<th>Agreem</th>
<th>Gene Name</th>
<th>Description</th>
<th>GO biological process</th>
<th>GO molecular function</th>
<th>GO cellular component</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.513</td>
<td>0.209</td>
<td>0.408</td>
<td>0.0187</td>
<td>100%</td>
<td>Aqp7</td>
<td>aquaporin 7</td>
<td>water transport</td>
<td>water channel activity</td>
<td>integral to plasma membrane</td>
</tr>
<tr>
<td>0.432</td>
<td>0.186</td>
<td>0.431</td>
<td>0.0089</td>
<td>100%</td>
<td>Bmp10</td>
<td>bone morphogenetic protein 10</td>
<td>skeletal development; pattern specification; TGFbeta receptor</td>
<td>cytokine; growth factor</td>
<td></td>
</tr>
<tr>
<td>0.700</td>
<td>0.335</td>
<td>0.478</td>
<td>0.0793</td>
<td>75%</td>
<td>Bmp15</td>
<td>bone morphogenetic protein 15</td>
<td>skeletal development; pattern specification; TGFbeta receptor</td>
<td>cytokine; growth factor</td>
<td></td>
</tr>
<tr>
<td>0.570</td>
<td>0.005</td>
<td>0.009</td>
<td>0.0000</td>
<td>100%</td>
<td>Bmp5</td>
<td>bone morphogenetic protein 5</td>
<td>skeletal development; pattern specification; TGFbeta receptor</td>
<td>cytokine; growth factor</td>
<td></td>
</tr>
<tr>
<td>0.693</td>
<td>0.143</td>
<td>0.207</td>
<td>0.0033</td>
<td>100%</td>
<td>Bmp7</td>
<td>bone morphogenetic protein 7</td>
<td>capping protein alpha 3</td>
<td>actin binding; F-actin capping</td>
<td></td>
</tr>
<tr>
<td>0.612</td>
<td>0.024</td>
<td>0.039</td>
<td>0.0000</td>
<td>100%</td>
<td>Cappa3</td>
<td>capping protein alpha 3</td>
<td>actin cytoskeleton organization and biogenesis</td>
<td>apoptosis</td>
<td></td>
</tr>
<tr>
<td>0.588</td>
<td>0.017</td>
<td>0.030</td>
<td>0.0000</td>
<td>100%</td>
<td>Casp8</td>
<td>Caspase 8</td>
<td>caspase activity</td>
<td>cytoplasm</td>
<td></td>
</tr>
<tr>
<td>0.679</td>
<td>0.070</td>
<td>0.103</td>
<td>0.0229</td>
<td>100%</td>
<td>Cone1</td>
<td>cyclin E1</td>
<td>mitosis; cell cycle; cytokinesis; regulation of cell cycle</td>
<td>cyclin-dependent protein kinase, regulator</td>
<td></td>
</tr>
<tr>
<td>0.800</td>
<td>0.064</td>
<td>0.079</td>
<td>0.0121</td>
<td>100%</td>
<td>Ccnf</td>
<td>cyclin F</td>
<td>mitosis; cell cycle; cytokinesis; regulation of cell cycle</td>
<td>cyclin-dependent protein kinase, regulator</td>
<td></td>
</tr>
<tr>
<td>0.543</td>
<td>0.178</td>
<td>0.328</td>
<td>0.0210</td>
<td>100%</td>
<td>Ccnf</td>
<td>cyclin G</td>
<td>mitosis; cell cycle; cytokinesis; regulation of cell cycle</td>
<td>cyclin-dependent protein kinase, regulator</td>
<td></td>
</tr>
<tr>
<td>0.684</td>
<td>0.139</td>
<td>0.203</td>
<td>0.0026</td>
<td>100%</td>
<td>Ccnf2</td>
<td>cyclin G2</td>
<td>cell cycle; cytokinesis; regulation of cell cycle</td>
<td>extracellular space</td>
<td></td>
</tr>
<tr>
<td>0.744</td>
<td>0.005</td>
<td>0.006</td>
<td>0.0000</td>
<td>100%</td>
<td>Cdd6;</td>
<td>Mus musculus gene for Macrolatin, complete cds</td>
<td>cell cycle; cytokinesis; protein amino acid phosphorylation</td>
<td>kinase; ATP binding; transferase; protein kinase</td>
<td></td>
</tr>
<tr>
<td>0.640</td>
<td>0.103</td>
<td>0.161</td>
<td>0.0385</td>
<td>100%</td>
<td>Cdk6</td>
<td>cyclin-dependent kinase 6</td>
<td>cell cycle; cytokinesis; protein amino acid phosphorylation</td>
<td>cell cycle; cytokinesis; protein amino acid phosphorylation</td>
<td></td>
</tr>
<tr>
<td>0.749</td>
<td>0.166</td>
<td>0.222</td>
<td>0.0138</td>
<td>100%</td>
<td>Cdkn2c</td>
<td>cyclin-dependent kinase</td>
<td>chromatin organizing structure</td>
<td>chromatin organizing structure</td>
<td></td>
</tr>
<tr>
<td>0.710</td>
<td>0.010</td>
<td>0.015</td>
<td>0.0000</td>
<td>100%</td>
<td>Chga</td>
<td>chromogranin A</td>
<td>cyclic nucleotide gated channel beta 3</td>
<td>vision; ion transport; cation transport; potassium transport electron transport</td>
<td>vision; ion transport; cation transport; potassium transport electron transport</td>
</tr>
<tr>
<td>0.557</td>
<td>0.268</td>
<td>0.482</td>
<td>0.0098</td>
<td>100%</td>
<td>Cngb3</td>
<td>cyclic nucleotide gated channel beta 3</td>
<td>cytochrome c oxidase, subunit IV</td>
<td>cytochrome c oxidase, subunit IV</td>
<td></td>
</tr>
<tr>
<td>0.745</td>
<td>0.227</td>
<td>0.304</td>
<td>0.0402</td>
<td>100%</td>
<td>Cox4a</td>
<td>cytochrome c oxidase</td>
<td>oxidoreductase; cytochrome c oxidase</td>
<td>oxidoreductase; cytochrome c oxidase</td>
<td></td>
</tr>
<tr>
<td>0.594</td>
<td>0.221</td>
<td>0.373</td>
<td>0.0065</td>
<td>100%</td>
<td>Cox5a</td>
<td>cytochrome c oxidase</td>
<td>electron transport</td>
<td>electron transport</td>
<td></td>
</tr>
<tr>
<td>Sh3bp3</td>
<td>SH3-domain binding protein 3</td>
<td>Insulin receptor signaling pathway</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------</td>
<td>----------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.535</td>
<td>0.011</td>
<td>0.021</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.351</td>
<td>0.009</td>
<td>0.017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.058</td>
<td>0.002</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.059</td>
<td>0.002</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.537</td>
<td>0.009</td>
<td>0.026</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.543</td>
<td>0.012</td>
<td>0.022</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.594</td>
<td>0.002</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.531</td>
<td>0.009</td>
<td>0.017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.535</td>
<td>0.011</td>
<td>0.021</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.521</td>
<td>0.020</td>
<td>0.039</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.477</td>
<td>0.183</td>
<td>0.383</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.750</td>
<td>0.122</td>
<td>0.162</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.685</td>
<td>0.152</td>
<td>0.222</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.635</td>
<td>0.291</td>
<td>0.458</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.645</td>
<td>0.174</td>
<td>0.270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.641</td>
<td>0.186</td>
<td>0.290</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.527</td>
<td>0.188</td>
<td>0.356</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.789</td>
<td>0.124</td>
<td>0.158</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.665</td>
<td>0.178</td>
<td>0.267</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.637</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.486</td>
<td>0.023</td>
<td>0.047</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.560</td>
<td>0.002</td>
<td>0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.428</td>
<td>0.002</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.652</td>
<td>0.036</td>
<td>0.055</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.778</td>
<td>0.216</td>
<td>0.277</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.500</td>
<td>0.012</td>
<td>0.029</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.609</td>
<td>0.198</td>
<td>0.325</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.576</td>
<td>0.150</td>
<td>0.261</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.677</td>
<td>0.018</td>
<td>0.026</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.691</td>
<td>0.026</td>
<td>0.037</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.704</td>
<td>0.029</td>
<td>0.039</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.542</td>
<td>0.004</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.538</td>
<td>0.165</td>
<td>0.306</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.720</td>
<td>0.172</td>
<td>0.239</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.747</td>
<td>0.256</td>
<td>0.342</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.652</td>
<td>0.036</td>
<td>0.055</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.778</td>
<td>0.216</td>
<td>0.277</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.531</td>
<td>0.009</td>
<td>0.017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.535</td>
<td>0.011</td>
<td>0.021</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(table 4A continued)

<table>
<thead>
<tr>
<th>Ave. Ratio</th>
<th>SD</th>
<th>CV</th>
<th>P-Value</th>
<th>Agreement</th>
<th>Gene Name</th>
<th>Description</th>
<th>GO biological process</th>
<th>GO molecular function</th>
<th>GO cellular component</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.513</td>
<td>0.012</td>
<td>0.033</td>
<td>0.0088</td>
<td>100%</td>
<td>Sort1</td>
<td>sortin 1</td>
<td>Endocytosis</td>
<td>neurotransin receptor activity, G-protein coupled receptors</td>
<td>cellular component</td>
</tr>
<tr>
<td>0.603</td>
<td>0.218</td>
<td>0.361</td>
<td>0.0012</td>
<td>83%</td>
<td>Spa1</td>
<td>alpha-spectrin 1, erythrocyte</td>
<td>rhodopsin-like receptor activity</td>
<td>G-protein coupled receptor protein signaling pathway</td>
<td>integral to membrane</td>
</tr>
<tr>
<td>0.439</td>
<td>0.010</td>
<td>0.023</td>
<td>0.0000</td>
<td>100%</td>
<td>Sprn2a</td>
<td>Mouse Musculus SPRN2A gene</td>
<td>regulation of cell shape</td>
<td>Structural constituent of cytoskeleton</td>
<td>extracellular space</td>
</tr>
<tr>
<td>0.506</td>
<td>0.017</td>
<td>0.033</td>
<td>0.0000</td>
<td>100%</td>
<td>S6k22c</td>
<td>serine/threonine kinase 22C</td>
<td>protein amino acid phosphorylation</td>
<td>extracellular process</td>
<td>cytoskeleton</td>
</tr>
<tr>
<td>0.635</td>
<td>0.001</td>
<td>0.002</td>
<td>0.0000</td>
<td>100%</td>
<td>Tbp</td>
<td>thrombopoietin</td>
<td>protein tyrosine kinase activity</td>
<td>extracellular process</td>
<td>cytoskeleton</td>
</tr>
<tr>
<td>0.714</td>
<td>0.030</td>
<td>0.042</td>
<td>0.0000</td>
<td>100%</td>
<td>Timeless</td>
<td>timeless homolog (Drosophila)</td>
<td>branching morphogenesis</td>
<td>DNA binding, protein binding</td>
<td>cytoskeleton</td>
</tr>
<tr>
<td>0.476</td>
<td>0.141</td>
<td>0.296</td>
<td>0.0011</td>
<td>100%</td>
<td>Tmod1</td>
<td>tropomodulin 1</td>
<td>filament organization</td>
<td>Peroxidase activity</td>
<td>cytoskeleton</td>
</tr>
<tr>
<td>0.556</td>
<td>0.348</td>
<td>0.626</td>
<td>0.0118</td>
<td>100%</td>
<td>Tnn2</td>
<td>tropomin I, skeletal, fast 2</td>
<td>filament organization</td>
<td>Hydrogen exporting ATPase activity</td>
<td>tropomin complex</td>
</tr>
<tr>
<td>0.669</td>
<td>0.310</td>
<td>0.463</td>
<td>0.0475</td>
<td>83%</td>
<td>Tpm2</td>
<td>tropomyosin 2, beta</td>
<td>muscle contraction</td>
<td>Structural constituent of cytoskeleton</td>
<td>muscle thin filament tropomyosin</td>
</tr>
</tbody>
</table>

Expression ratios are calculated from n = 6 replicated arrays. Normalization scheme and data filtration protocol are described in the Material and Methods section.

Table 4B. Selected up-regulated genes from the E-Tmod$^{+/−}$ cardiac tissue

<table>
<thead>
<tr>
<th>Ave. Ratio</th>
<th>SD</th>
<th>CV</th>
<th>P-Value</th>
<th>Agreement</th>
<th>Gene Name</th>
<th>Description</th>
<th>GO biological process</th>
<th>GO molecular function</th>
<th>GO cellular component</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.325</td>
<td>0.248</td>
<td>0.187</td>
<td>0.033</td>
<td>100%</td>
<td>Actg</td>
<td>actin, gamma, cytoplasmic</td>
<td>muscle development</td>
<td>motor activity</td>
<td>actin cytoskeleton</td>
</tr>
<tr>
<td>1.578</td>
<td>0.695</td>
<td>0.440</td>
<td>0.097</td>
<td>83%</td>
<td>Actn4</td>
<td>alpha-actinin alpha 4</td>
<td>cell motility, actin filament bundle formation</td>
<td>protein N-terminus binding</td>
<td>cytoplasm, cortical cytoskeleton</td>
</tr>
<tr>
<td>1.416</td>
<td>0.476</td>
<td>0.336</td>
<td>0.085</td>
<td>83%</td>
<td>Cappb1</td>
<td>capping protein beta 1</td>
<td>actin filament bundle formation</td>
<td>actin binding; F-actin capping</td>
<td>actin capping protein</td>
</tr>
<tr>
<td>1.966</td>
<td>0.901</td>
<td>0.458</td>
<td>0.047</td>
<td>100%</td>
<td>Dnc1c1</td>
<td>dynamin, cytoplasmic, light chain 1 dynamin 2</td>
<td>actin filament organization</td>
<td>microtubule motor activity</td>
<td>cytoplasmic dynein complex</td>
</tr>
<tr>
<td>1.692</td>
<td>0.364</td>
<td>0.215</td>
<td>0.006</td>
<td>100%</td>
<td>Dnm2</td>
<td>dynamin 2</td>
<td>protein binding</td>
<td>receptor mediated endocytosis</td>
<td>clathrin-coated endocytic vesicle</td>
</tr>
<tr>
<td>1.910</td>
<td>0.009</td>
<td>0.005</td>
<td>0.000</td>
<td>100%</td>
<td>Melv</td>
<td>Mediterranean fever</td>
<td>unknown</td>
<td>actin binding/zinc ion binding</td>
<td>intracellular</td>
</tr>
<tr>
<td>1.520</td>
<td>0.576</td>
<td>0.379</td>
<td>0.078</td>
<td>66%</td>
<td>Myh7</td>
<td>myosin, heavy polypeptide 7, cardiac muscle, beta</td>
<td>muscle development; cytoskeleton organization and biogenesis</td>
<td>motor; ATP binding; calmodulin binding</td>
<td>myosin; cytoskeletal thin filament</td>
</tr>
<tr>
<td>1.902</td>
<td>0.784</td>
<td>0.412</td>
<td>0.037</td>
<td>83%</td>
<td>Mym10</td>
<td>myosin X</td>
<td>cytoskeleton organization and biogenesis</td>
<td>motor; ATP binding; calmodulin binding</td>
<td>myosin; cytoskeleton</td>
</tr>
<tr>
<td>1.657</td>
<td>0.462</td>
<td>0.279</td>
<td>0.090</td>
<td>100%</td>
<td>Mym11</td>
<td>myosin IF</td>
<td>cytoskeleton organization and biogenesis</td>
<td>motor; ATP binding; calmodulin binding</td>
<td>myosin; cytoskeleton</td>
</tr>
<tr>
<td>4.644</td>
<td>0.080</td>
<td>0.017</td>
<td>0.010</td>
<td>100%</td>
<td>Rasa3</td>
<td>RAS p21 protein activator 3</td>
<td>signal transduction</td>
<td>GTPase activator</td>
<td>plasma membrane</td>
</tr>
</tbody>
</table>

Expression ratios are calculated from n = 6 replicated arrays. Normalization scheme and data filtration protocol are described in the Material and Methods section.

F. Clusters of differentially expressed genes analyzed by GeneSet Enrichment Analysis (GSEA)

knowledge-based GSEA analysis serves to reveal the up-and down-regulated functional gene clusters, rather than individual differentially expressed genes, from the microarray data. In E-Tmod$^{+/−}$ cardiac tissues, “cell cycle” is the top-ranked down-regulated gene cluster, while “inhibition of cell cycle”, “signal transduction”, “immune response” and “lipid/amino acid transport” are the highest-ranked up-
regulated gene groups (Fig. 11). Comparison between the top three clusters by the conventional and GSEA methods (table 5A) with the member of the affiliated genes (table 5B) are listed for both the up- and down-regulated categories.

Figure 11. Percentage distribution of the down-regulated (A) and up-regulated (B) gene-clusters revealed by non-parametric GeneSet Enrichment Analysis (GSEA), which identifies differentially expressed gene-groups from the non-differentially expressed ones by clustering with dual (expression and functional) parameters.

Table 5. Summary of DE genes from top ranked clusters

(A) Top-ranked DE gene-clusters identified by the conventional vs. GSEA clustering methods, and the list of the DE genes (B) reported by both clustering analysis.
G. Validation of the microarray reported DE genes with qPCR assay

Twelve genes that represent three of the largest DE clusters were selected for qPCR testing to confirm their microarray expression pattern (Table 6). These genes included the cyclins required for S phase (CDK6 & cyclin G1), G1 phase (cyclin G2) and G2 phase (cyclin F), as well as myogenic factor (BMP 5), and genes encoding cardiac thin-filament component proteins (α-cardiac actin), troponymosin 2 (TPM2), and Rasa3, an inhibitor for Ras induced proliferation pathway. Ten out the twelve genes assayed by qPCR confirmed the down-regulated expression reported in the microarray experiments. Cyclin E1 and Cyclin D2 are the only two genes showing lack of agreement between array and qPCR results. While both are shown to be down-regulated in microarray data, Cyclin E1 is found to be up-regulated and Cyclin D2 showed no change in the qPCR assay.

Table 6. Selected DE genes detected by microarray and qPCR.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Microarray</th>
<th>qPCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actc1</td>
<td>0.67</td>
<td>0.33</td>
</tr>
<tr>
<td>Actn4</td>
<td>1.58</td>
<td>1.62</td>
</tr>
<tr>
<td>BMP5</td>
<td>0.57</td>
<td>0.38</td>
</tr>
<tr>
<td>Cappb1</td>
<td>1.42</td>
<td>1.78</td>
</tr>
<tr>
<td>CDK5</td>
<td>0.70</td>
<td>0.71</td>
</tr>
<tr>
<td>CDK6</td>
<td>0.64</td>
<td>0.50</td>
</tr>
<tr>
<td>Cyclin E2</td>
<td>0.52</td>
<td>0.70</td>
</tr>
<tr>
<td>Cyclin D1</td>
<td>0.34</td>
<td>0.19</td>
</tr>
<tr>
<td>Cyclin D2</td>
<td>1.25</td>
<td>0.97</td>
</tr>
<tr>
<td>Cyclin E1</td>
<td>0.68</td>
<td>1.22</td>
</tr>
<tr>
<td>Cyclin F</td>
<td>0.79</td>
<td>0.28</td>
</tr>
<tr>
<td>Cyclin G1</td>
<td>0.54</td>
<td>0.71</td>
</tr>
<tr>
<td>Dynein</td>
<td>1.96</td>
<td>2.72</td>
</tr>
<tr>
<td>E-Tmod</td>
<td>0.62</td>
<td>0.51</td>
</tr>
<tr>
<td>Mras</td>
<td>0.39</td>
<td>0.62</td>
</tr>
<tr>
<td>Rasa3</td>
<td>4.64</td>
<td>3.69</td>
</tr>
<tr>
<td>TPM 2</td>
<td>0.67</td>
<td>0.41</td>
</tr>
<tr>
<td>U-Tmod</td>
<td>2.93</td>
<td>2.18</td>
</tr>
</tbody>
</table>

Expression ratios of both columns were calculated from expression level normalized against the wildtype controls.
H. Differentially expressed genes in E-Tmod$^{-/-}$ and E-Tmod$^{+/-}$ cardiac tissues

Comparison of the DE genes from cardiac tissues of E-Tmod$^{-/-}$ and E-Tmod$^{+/-}$ mice revealed a similar pattern of regulation (table 7). Genes from various functional clusters, including cyclin, myogentc and members of actomyosin family are down-regulated, while Rasa3, an inhibitor for cellular proliferative signalings, is up-regulated. Although mature hearts taken from the E-Tmod$^{+/-}$ samples cannot be considered equivalent to the less developed embryonic hearts, the common regulation pattern shared by E-Tmod$^{+/-}$ and E-Tmod$^{-/-}$ mice may provide clues as to whether differential expression of particular genes may contribute to the cardiac phenotype observed.

Table 7. Differentially expressed genes in E-Tmod$^{+/-}$ and E-Tmod$^{-/-}$ hearts.

<table>
<thead>
<tr>
<th>Gene</th>
<th>E-Tmod$^{+/-}$ (adult)</th>
<th>E-Tmod$^{-/-}$ (embryo)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-Tmod</td>
<td>0.51</td>
<td>0.00</td>
</tr>
<tr>
<td>ATCT</td>
<td>0.33</td>
<td>0.59</td>
</tr>
<tr>
<td>TPM 1</td>
<td>0.64</td>
<td>0.00</td>
</tr>
<tr>
<td>TPM 2</td>
<td>0.41</td>
<td>0.91</td>
</tr>
<tr>
<td>TNNi</td>
<td>0.56</td>
<td>0.50</td>
</tr>
<tr>
<td>BMP-5</td>
<td>0.38</td>
<td>0.70</td>
</tr>
<tr>
<td>Cyclin-D1</td>
<td>0.34</td>
<td>0.63</td>
</tr>
<tr>
<td>RASA3</td>
<td>3.69</td>
<td>3.17</td>
</tr>
<tr>
<td>U-Tmod</td>
<td>2.18</td>
<td>2.24</td>
</tr>
</tbody>
</table>

The E-Tmod$^{+/-}$ samples (n = 3) were extracted from male mice at 20 weeks of age, while E-Tmod$^{-/-}$ hearts (n = 3) were collected from embryos at gestational day 9.5. Expression ratios from both groups were calculated from the expression level normalized against the wildtype controls.

Discussion

In the present study, we have established the expressions of E-Tmod deficient mice of both homozygous and heterozygous types. Although the E-Tmod$^{-/-}$ mice exhibited more severe phenotypes and are embryonically lethal, the cardiac tissues of the Tmod$^{+/-}$ mice revealed a panel of differentially expressed genes that is
similar to their null KO counterparts. Specifically, the thin filament-associated genes, such as ACTC and TPMs that encode cardiac abundant actin and tropomyosins, are down-regulated significantly in both $E-Tmod^{-/-}$ and $Tmod^{+/+}$ mice.

Surveying the tropomodulin homologs in $E-Tmod^{+/+}$ and $E-Tmod^{-/-}$ embryos revealed a significant induction of the ubiquitous $Tmod$ ($U-Tmod$) levels relative to the wildtype samples. Although the indication of compensatory of $U-Tmod$ for $E-Tmod$ deficiency is demonstrated at the transcriptional level, it is clear that this response is not significant enough to rescue the phenotype of the $E-Tmod^{-/-}$ mice. On the other hand, survival of $E-Tmod^{+/+}$ mice indicates that a single copy of $E-Tmod$ gene is sufficient to sustain the normal cardiac functions. This outcome demonstrated the irreplacibility of the $E-Tmod$ and the robustness of its single allele.

In cardiac myofibrils, tropomyosin (TM) helical chains are formed by heterodimers (Clark et al., 2002) encoded by TPM1 and TPM2, with the $\alpha$-TM representing the more abundant specie (Wolska et al., 2003), and cardiac actin representing the major (80%) actin isoform in the adult heart (Herman et al., 1993). Absence of TPM1 in the E-Tmod null embryo and lowered amount of TPM2 in $E-Tmod^{+/+}$ heart, compounded by the reduction of ACTC1 expression in the both $E-Tmod^{-/-}$ and $T-tmod^{+/+}$ cardiac tissues (by 33% and 47% respectively), revealed a common trend of multiple thin filament component down-regulation in the E-Tmod deficient hearts.

It has been reported that genes mediating actin-filament turnover (Arp2/3, profilin, gesolin, and capping proteins) are tightly co-regulated (Vlacuou et al.,
2005) during growth and cytoskeletal reorganization. The knockout of a barbed-end actin capping protein, β-adducin, resulted in altered expressions of thin-filament components in mouse erythrocytes (Porro et al., 2004). In particular, the compensatory up-regulation of E-Tmod and capZ, accompanied by a significant down-regulation of actin and tropomyosin, clearly suggested that capping protein, such as β-adducin, can regulate the expressions of other components in the actin thin filament.

Utilizing the GeneSet Enrichment Analysis (GSEA) algorithm in combination with the conventional clustering technique, the microarray expression profiling revealed multiple members of the cell cycle regulatory genes being down-regulated in the Tmod+/− hearts. Notably, cyclin D1 (33%), E1 (68%), E2 (52%) are known to regulate the cell size of the terminally differentiated cardiomyocytes in post-natal rodents (Busk et al., 2002, Fortuño et al., 2001). The down-regulation of the cyclin-cluster expressions implicates a reduced capability to increase in cell-size, which provides an underlying mechanism for the cardiac underdevelopment observed in the E-Tmod deficient mice.

In addition to a panel of cyclins differentially expressed in E-Tmod whole hearts, myogenic growth factors MyoD1 (48% ± 17%) and a group of BMPs (BMP5 = 57%, BMP7 = 69%, and BMP10 = 43%) are also notably down-regulated. While BMPs are important in signaling through TAK1 and SMADs that are necessary for the differentiation of cardiomyocytes (Monzen et al., 1999). MyoD1 is a crucial determinant in the progression of dystrophy-associated cardiomyopathy (Megeney et al., 1999), and it is known as a transcriptional regulator governing TM expression (Lees-Miller and Helfman, 1991) in striated
muscles. Down-regulation of these myogenic markers in \( E-Tmod^{+/−} \) cardiac tissue suggested a retardation of differentiation activity crucial to growth and development.

We noted two of the 13 genes being validated with qPCR show discrepancies in their expression with the array result. This phenomenon may be explained by the variability of the biological samples (different mouse samples were used). Though similarly matched in age and weight, variations in gene expression of \textit{in vivo} mouse specimen are often observed (Etienne \textit{et al.}, 2004). Furthermore, primers for qPCR amplification and oligo probes on the microarray do not cover the same length and segment of sequences in these genes, which could lead to different base-pairing efficiencies and the quantification results. From the data analysis perspective, qPCR expression ratios were normalized to a single housekeeping gene (GAPDH), but microarray expressions were normalized to a population of pseudo-housekeeping genes (median distribution). The difference in normalizaiton and quantification techniques may also contribute to the differences observed between the two different expression assay results.

Defects and altered expression in sarcomeric components have been associated with numerous cardiac phenotypes (Marian \textit{et al.}, 1995, Mogensen \textit{et al.}, 1999 and Muthuchamy \textit{et al.}, 1998). We have previously shown that E-Tmod null (homozygous knockout) mice have severe cardiac phenotypes of ventricular underdevelopment at early embryonic stage (Chu \textit{et al.}, 2003). The present study has identified, in both \( E-Tmod^{−/−} \) and \( Tmod^{+/−} \) mice, several common differentially-expressed genes potentially contribute to the observed cardiac abnormalities. Through the profiling and clustering of differentially expressed genes in the E-Tmod deficient mice based on functional classifications, we have
shown by microarray study the association between specific genes and the potentially affected phenotypes in a genetic knockout model.
Chapter III

Hypoxia-Induced Pressure Overload of E-Tmod deficient Mouse Heart

Abstract

Tropomodulin is an actin-capping protein associated with both sarcomeric and cytoskeletal actin filaments in cardiac muscle. Homozygous knockout (KO) of erythrocyte tropomodulin (E-Tmod) is embryonically lethal, but heterozygous knockout (+/-) mice survive. Heterozygous E-Tmod KO resulted in smaller right ventricle (RV) cavities and free walls compared to the wild type. To investigate the effect of heterozygous E-Tmod KO on mouse cardiac function and remodeling, mice (n = 7 for E-Tmod^{+/−} and n = 6 for E-Tmod^{+/+}) were subjected to 5 weeks of hypoxia to increase the pressure loading on the RV via pulmonary hypertension. The effect of loading was determined by measuring the volume (free wall and cavity) of RV anatomical features, and mechanical function during the cardiac cycle. In the wild type, although the pressure loading had no significant effects on RV cavity or free wall volume, geometrical measurements suggest a tissue redistribution. Under equal loading conditions, E-Tmod^{+/−} mice exhibited a significant increase in volume for both RV cavity and free wall. RV epicardial function showed an increase in surface area strain at peak systole for hypoxic knockout mice. Thus it appears the heterozygous KO of E-Tmod affects RV volume and function under normal and pressure loading conditions.
Introduction

Altered E-Tmod expression has been shown to cause significant abnormalities in various animal models. *In vitro* overexpression of E-Tmod in both rat and chick cardiomyocytes led to shortening of the thin filaments and the disorganization of sarcomeres, resulting in myofibril degeneration (Littlefield *et al.*, 2001). Likewise, mouse hearts with *in vivo* overexpression of E-Tmod showed disrupted sarcomere organization with shortened thin filaments, leading to myofibril degeneration and dilated cardiomyopathy (Sussman *et al.*, 1998).

Absence of both E-Tmod alleles (E-Tmod$^{-/-}$) results in lethality of the mouse embryo at approximately E10. Phenotypically, these E-Tmod$^{-/-}$ embryos show retarded heart growth, disrupted vessel development, and restricted erythrogenesis in the yolk sac (Chu *et al.*, 2003), demonstrating a profound importance of E-Tmod in the development of vital cardiovascular and possibly other systems. In view of the primary and early lethal defect in the cardiac tissues of the E-Tmod$^{-/-}$, we investigated the adaptive response and cardiovascular functions of E-Tmod$^{+/-}$ mice under normal and hypoxia conditions.

We hypothesize that E-Tmod may play dual roles in cardiac myocytes: (1) it may directly affect the force generation capacity of sarcomeres by modulating sarcomeric actin, and (2) it may affect the actin cytoskeleton and hence the role of mechanotransduction load regulation. In order to test these possibilities, the present study examined local contractile function in the right ventricles of E-Tmod$^{+/-}$ mice compared to control, and overloaded the right ventricles with pulmonary hypertension due to hypoxia and examined the role of E-Tmod in the tissue hypertrophy response.
Materials and Methods

A. Hypoxia conditioning of E-Tmod

Control and E-Tmod mice were subjected to hypoxia to examine the role of E-Tmod in ventricular remodeling. A transparent chamber with dimensions of 20” x 18” x 30’ (W x D x H) was constructed with thermoplastic resin walls. An oxygen monitor (ProOX model 110, Reming Bioinstruments Co., Redfield, NY) and valve regulator (CP701-125-580-AG, Harris Specialty Gas, Gainesville, GA) were coupled to a nitrogen tank to provide gas to the hypoxia chamber to reduce the oxygen content via feedback control, maintaining the oxygen level at 11% within the enclosed environment.

B. Collection of hearts from hypoxia-treated mice

At the end of the hypoxia conditioning period, the mice were anesthetized using ketamine (100 mg/kg, Fort Dobe Animal Health, Fort Dobe, IW)-xylazine (8 mg/kg, Vedco Inc., St. Joseph, MO) and placed on a Rodent Ventilator Model 683 (Havard Apparatus, South Natick, MA). The chest was opened, and the left ventricle was injected with 0.5 to 1.0 ml of an arresting solution containing 4.0 g/L NaCl, 2.2 g/L KCl, 1.0 g/L NaHCO3, 2.0 g/L dextrose, and 10 mg/L heparin and 3.0 2,3-butanedione monoxime (BDM). The arrested heart was excised from the mouse, weighted, and a 30-gauge needle was inserted into the ascending aorta and out of the cardiac apex in order to define the long axis of the heart for subsequent sectioning. The heart was then frozen in this unloaded state (zero cavity pressure) in a mixture of dry ice and methyl butane, then frozen for storage.
C. Functional analysis of hypoxia-treated hearts

To measure local function, hearts were exposed (from the same set of mice in section B), a COHU solid-state camera was focused on the RV, and a spray of powdered titanium dioxide was painted onto the surface of the heart (Figure 12B). Video data sets and ECG (Gould ECG/BioTach monitor) were recorded for each mouse (Figure 12A). The material surface markers were mapped onto a Cartesian plane using ImageJ software. Data was processed for two-dimensional homogeneous strain at the central portion of the right ventricular free wall. Systolic function was determined as change in surface area at peak-systole relative to end-diastole of each right ventricle. End-diastole was taken as the ECG R peak.

![Figure 12. ECG capture (A) and synchronization with video (B) of titanium dioxide labeled mouse heart. Data were processed for relative strain and change in RV surface area during systole.](image)

D. Sectioning and staining of whole heart tissues

A Leica CM3050 Cryostat was used to cut 20-mm thick frozen sections every 400 mm throughout short axis along each heart. These sections were taken perpendicular to the long axis of the heart defined by the aorta-apex line. After staining with a solution of picric acid (100 ml) and Sirius red (0.1 g), a collagen specific stain, the sections were subjected to microscopic (Nikon OptiPhot-2 with a 5x objective) observation and digitized with a digital camera (Nikon CoolPix 990),
The NIH imaging software (http://rsb.info.nih.gov/nih-image/about.html) were used for contrast adjustment and segmentation, and to calculate the right ventricular chamber area and the right ventricular free-wall area from each slice. Simpson’s trapezoid method of discs was employed for calculating the right ventricular chamber volume in each heart by compiling the sectional areas.

Results

A. Heart weight in normoxia and hypoxia

Prior of performing histological analysis, the whole hearts collected from individual mice of each group were weighed and normalized by their corresponding body weights. The average total heart weight, which includes all chambers (atrium and ventricle), did not show any significant difference between the control and hypoxia groups nor between the wildtype and E-Tmod$^{+/−}$ mice (Fig. 13).

![Figure 13](image)

Figure 13. Normalized heart weight of mice without (white columns) and with (gray columns) hypoxia treatment. The numerical values of the heart/body weight ratios of each group are indicated in the bars, and numbers of wildtype (+/+) and E-Tmod heterozygous knockout (+/-) animals are given in parenthesis.
B. Hematocrit response of E-Tmod mice to hypoxia

Mice from both E-Tmod<sup>−/+</sup> mice and E-Tmod<sup>++/+</sup> groups were subjected to either hypoxic stress (11.0% O<sub>2</sub>) in a hypoxia chamber, or kept in the same room at normal atmospheric composition (20.9% O<sub>2</sub>). After 5 weeks of hypoxia treatment, both E-Tmod<sup>−/+</sup> and E-Tmod<sup>++/+</sup> groups showed elevated hematocrit, but there are no statistically significant difference between the two genotypes (Fig. 14).

![Graph showing hematocrit levels before and after hypoxia treatment for E-Tmod<sup>−/+</sup> and E-Tmod<sup>++/+</sup> mice.](image)

Figure 14. Hematocrit levels of E-Tmod<sup>−/+</sup> and E-Tmod<sup>++/+</sup> mice under normoxia and after the hypoxia treatment. The numerical values of the hematocrit of each group are indicated in the bars, and numbers of wildtype (+/+), and E-Tmod heterozygous knockout (+/-) animals are given in parenthesis. Numbers of animals are labeled in parenthesis * indicates a statistically significant difference (P < 0.01) between the hypoxic and normoxic groups.

C. Cardiac response to hypoxia

Quantitative analysis of the whole-heart sections (Fig. 15) showed that the normoxic E-Tmod<sup>−/+</sup> mice have smaller RV cavity and free wall volume than all other groups (P < 0.01, Fig. 16A). 5 weeks of hypoxia treatment did not have any effect in the RV cavity and free wall volume in E-Tmod<sup>++/+</sup> mice (Fig. 16B), but caused increases in the E-Tmod<sup>−/+</sup> mice to become not significantly different from the values obtained from the E-Tmod<sup>++/+</sup> group.
Figure 15. Samples of $E-Tmod^{-/-}$ and $E-Tmod^{+/+}$ whole heart sections under normoxic and hypoxic conditions. Normalized RV cavity and RV wall volume were complied according to procedures described in the Materials and Methods section.

Figure 16. (A) RV cavity volume and (B) Ratio of RV/LV wall volume, of $E-Tmod^{-/-}$ and $E-Tmod^{+/+}$, are compared with hypoxic and normoxic conditions. Statistical significance ($P < 0.05$, * ) was revealed in the $E-Tmod^{-/-}$ mice.
D. Local remodeling of RV in E-Tmod\(^{+/+}\) after hypoxia treatment

Although wild type mice demonstrated little global response to the hypoxia treatment, sections of the free wall exhibited significant hypertrophy (see Fig. 17). 2D geometrical modeling (Fig. 17) suggests that the arrowed region (Fig. 17C) of the right ventricular free wall was hypertrophied. Such local adaptations to hypoxia may occur in areas where the stimulus to hypertrophy is greatest.

![Figure 17. Right ventricle free wall for structural study of wild type mice under normoxic and Hypoxic conditions. (A) location of modeling; (B) RV free wall volume for normoxic wild type and (C) RV free wall volume for hypoxic wild type.](image)

E. Functional analysis

In order to study the effect of E-Tmod deficiency on contractile function, local deformation was investigated on the RV surface under both normal and hypoxic conditions. The pressure loading due to hypoxia caused the maximum relative surface area during the cardiac cycle to change by 20.4% in the E-Tmod\(^{+/+}\) mice. In comparison, the maximum surface area change in E-Tmod\(^{+/+}\) under hypoxia is only 14.7%. This finding shows that E-Tmod\(^{+/+}\) mouse hearts subjected to hypoxia have increased contractile function of the RV surface relative to the wildtype mice in either conditions.
Discussion

In the current study, we report for the first time that the RV cavity and wall volumes of the E-Tmod\(^{+/−}\) mice are significantly smaller than their wildtype. This finding is in agreement with results from the gene expression profiling (Chapter 2) that showed E-Tmod\(^{+/−}\) mice have lower amounts of cardiac thin filament components, as well as reduced levels of growth and differentiation related genes, and in accordance with fewer thin filaments found in the cardiomyocytes derived from E-Tmod\(^{−/−}\) ES cells (Ono, 2005). Although no disease phenotype is observed in the E-Tmod\(^{+/−}\) mice, probably because the smaller RV chamber and thinner wall can deal with the low pulmonary pressure (~20 mmHg) (Steudel et al., 1998) of E-Tmod\(^{+/−}\) mice. The smaller cavity and free wall volume in E-Tmod\(^{+/−}\) mice compared to the wild type (Fig. 16) is consistent with previous reports that E-Tmod plays an integral role in sarcomere structure. With reduced E-Tmod, actin filaments would grow unregulated and disorganized, which could interfere with proper sarcomeric protein interactions and decrease strain. This
could lead to abnormal growth of the ventricle including a lack in ability to adapt to growth cues.

Various types of cardiac remodeling in response to stress are known to involve fibers of both thin and thick filaments of the cardiomyocytes (Martin et al., 1996). With hypoxia treatment, the E-Tmod group showed restoration of RV cavity and wall volume to the normoxic sizes. Although RV cavity and free wall volume of E-Tmod\(^{+/+}\) mice did not change in respond to the level of the hypoxia given, these animals did show a significant elevation of the hematocrit level. In E-Tmod\(^{+/−}\) group, the significant increases of RV cavity in size and thickening (23.8% increase) of free wall thickness reflect cardiac tissue remodeling in respond to the hypoxic stress. Such structural changes have been shown to result from the increased pulmonary resistance induced by hypoxia stress (Steudel et al., 1998 & Fan et al., 2005), and this would be compounded by an increase in blood viscosity result from the higher hematocrit level. These effects are likely to stimulate the E-Tmod\(^{+/−}\) RV to undergo hypertrophic remodeling in order to increase thier cardiac output needed for the extra physiological demand.

Although the hypoxia conditioning in our setup did not impose a high enough loading to contribute to significant global hypertrophy in wild type mice, the stress was able to induce local RV remodeling (Fig.17) and significant elevation of hematocrits, indicating the hypoxic treatment is effective. The fact that hypoxia-treated E-Tmod\(^{+/−}\) mice ended up with comparable, but not larger RV cavity and wall volumes relative to the wildtypes suggests that this ventricular size is adequate for the degree of stress imposed.
The functional study on RV strain during systole shows the $E-T_{mod}^{+/−}$ mice have increased contractile function. Specifically, knockout mice under the increased loading conditions exhibited an 11.9% increase in surface area during systole as compared to wild type control, and a 9.9% increase compared to the wild type hypoxia group. In other words, wild type mice under either condition exhibited smaller change (less strain) in RV surface area compared to knockouts under adverse loading conditions. This finding is in agreement with our volume study which showed a smaller RV dimension in $E-T_{mod}^{+/−}$ under normoxia condition but increased more substantially after the hypoxia treatment compared to the wild type.

The current $in vivo$ study showed that heterozygous E-Tmod knockout mice have intact mechanotransduction pathways and retain the ability to undergo hypertrophy and remodeling under stress. These growth mechanisms may even be enhanced, as demonstrated by the functional study that suggests an increase in adaptive remodeling that compensates for the loss of the structural proteins, and this may explain the absence of disease phenotype in $E-T_{mod}^{+/−}$ mice.
Chapter IV

Mechanical Regulation of E-Tmod in Cardiomyocytes

Abstract

We hypothesize that E-Tmod, as a cytoskeletal component and sarcomeric protein plays a role in mechanically induced myocyte remodeling. Furthermore, we postulate that spliced isoforms of E-Tmod with distinct epitopes would lead to differential response to external mechanical stimuli. To test the above hypothesis, we cloned and characterized the E-Tmod splice variants in the neonate rat cardiomyocytes (NRC), and assessed their reactions to shear and stretch. Also investigated were the expressions of cardiac actin and α-tropomyosin, with the goal of elucidating the role of these mechanical stimuli (shear and stretch) on cardiac phenotype. Mechanical strain and shear also elevated E-Tmod in both wildtype and E-Tmod knockdown myocytes. These results demonstrated that upregulation of myofibril E-Tmod is simultaneous with the hypertrophic process and partial reduction of the E-Tmod (both E-Tmod +/- and siRNA treated cardiomyocytes) does not affect the ability of the cardiomyocytes to undergo stretch-induced adaptive hypertrophy.

Introduction

The heart is a dynamic organ that exhibits initial contractions at the 3-somite stage (E 8.25 in mice) and beats coordinately by 4-somite stage (E 9.25 in mice), leading to vital unidirectional blood flow during early embryonic development (Nishii & Shibata, 2006). Components of cardiac tissue therefore are subjected to constant homodynamic forces and their development and function are likely regulated by these forces. Embryonic mouse cardiomyocytes are known to
respond to mechanical strain with proliferation (Miller, 2000), while neonate rat cardiomyocytes (NRC) can undergo hypertrophy (Cooper et al., 1985 & Kira et al., 1984) and myofibril alignment (Chedrawy et al. 2002) when exposed to stretching. Transverse shearing between myocardial sheets is correlated with wall thickening (LeGrice et al., 1995) and can affect ectopic pacing (Lorenzen-Schmidt et al., 2006). These findings suggest that mechanotransduction of various type of loads can play a significant role in controlling cardiomyocyte function and homeostasis.

As most other cell types responsive to mechanostimulations, cardiomyocytes are believed to transduce external stimuli via a cytoskeleton that connects intracellular structures with the extracellular matrix (review article by Samarel, 2005). Actin-associated cytoskeletal proteins, therefore, are implicated to play pivotal roles in signal transduction and the remodeling of the cytoarchitecture when subjected to shear stress and mechanical stretch.

Various types of cardiac remodeling in response to mechanical stimuli involve fibers of both thin and thick filaments in cardiomyocytes (Martin et al. 1996). Such responses include cytoskeletal re-organization (Kawamura et al., 2003), changes in myofibril alignment (Kada et al., 1999), and alteration of cell size (Sadoshima and Izumo, 1993). Since E-Tmod is known to regulate the actin dynamics in various types of cells, investigation of the E-Tmod regulation in cellular remodeling in response to mechanical stimuli can provide novel information on its role in mechano-sensing and transduction.
Materials and Methods

A. Cloning of E-Tmod cDNA from rat cardiomyocytes

Total RNA from neonate rat cardiomyocytes was extracted with Trizol reagent (GIBCO-BRL, Life Technologies, Gaithersburg, MD) according to the manufacturer's protocol. Corresponding cDNA was synthesized by using Superscript reverse transcriptase (GIBCO-BRL) primed with oligo-dT. Resulting cDNA was used as a template for PCR with E-Tmod oligonucleotide upstream primer located within exon 0 sequence 5’-CGCCAGGGAACTCACCAG-3’ and the downstream primer 5’-TGTTCTACCGATAATGTGTTC-3’ that resides in exon 9. The resulting PCR products of the expected lengths (~1.1 Kb and 0.85 Kb) were subcloned into PCR 2.1 TOPO vector (Invitrogen, Carlsbad, CA) and sequenced for further analysis.

B. Protein assay and western blot

Ten mg of cardiomyocytes lysate was separated by SDS-PAGE on a 10% gel (pH 8.8) and transferred onto a nitrocellulose membrane. The membrane was incubated with an anti-E-Tmod antiserum (pAb, 1:100) (Sung and Lin, 1994) or anti-E-Tmod monoclonal antibody (G131-204 mAb, 1:150), followed by an HRP (horseradish peroxidase) conjugated goat-anti-rabbit IgG secondary antibody (1:1500) with color development or HRP conjugated goat-anti-mouse IgG (1:3000) with chemiluminescence.

C. Neonate rat cardiomyocyte culture

Primary cultured myocytes were prepared from ventricles of neonatal (within 3 days old) Spargue-Dawley rats using a commercial isolation kit (Cellutron Life Technology, Highland Park, NJ). Briefly, 30 neonate rats were decapitated, their
hearts quickly excised and trimmed to remove arterial tissue. The ventricles were gently stirred for 20 min in digestion buffer at 37°C. The cell suspension was centrifuged at 1200 rpm for 1 min, the supernatant discarded, and the pellet cells resuspended in buffer solution. This cycle was repeated a total of 5 times or until all of the tissue was digested. The cells were then pre-plated on uncoated plates for 1-2 hours at 37°C in order to reduce the contamination of cardiac fibroblasts (Gross et al., 1968). The cardiomyocytes were collected and plated on microscope slides coated with SureCoat at a plating density of 8 x 10^4 cells per cm², and cultured in plating medium contacting 15% horse serum (Gibco BRL, Paisley, UK). Culture medium was exchanged after 48 hours to standard maintenance medium containing 6% serum. The culture was allowed to incubate for 7 days on average, after which the cells had spread to a monolayer and were beating spontaneously at a stable rate.

**D. E-Tmod knockdown with siRNA**

Small interfering RNA targeting exon 2 of E-Tmod gene was designed (the segment which is identical between mouse and rat) using web-based siRNA designer (siRNA target finder, Ambion, Inc. Foster City, CA). The target sequence is 5’-GGACCGAGAAGATCTGGTT-3’. Sense and anti-sense siRNA were synthesized and mixed to form SiRNA duplex, which was transfected to cells using the lipid-based agent siRORT NeoFX (Ambio, Inc. Frellstedt, Germany) following the manufacturer’s instruction.

**E. Application of laminar shear**

A flow system (Fig. 19) was used to impose shear stress on cultured cardiomyocytes similar to that described previously (Lorenzen-Schmidt et al.
In brief, the glass slide with attached myocytes was assembled into a parallel plate flow chamber apparatus with dimensions (width x length x height) of 1.38 cm X 7.62 cm X 127 µm. The height of the flow chamber was set by placing a silastic gasket between the slide and the upper lid of the chamber. Input and output ports allowed the medium to be introduced as laminar flow over the entire chamber. Myocytes were subjected to a laminar shear stress of 12 dyn/cm² by using a mechanical syringe infusion pump (Harvard Apparatus, Holliston, MA). The flow system was kept at 37°C and equilibrated with 95% humidified air and 5% CO₂. The same batch of NRCs was used for three independent shear experiments.

Figure 19. Setup of shear flow system driven by a syringe pump. Laminar shear of 12 dyn/cm² was applied to the cardiomyocytes cultured on a coated glass slide. The system was kept in a 37°C hood and equilibrated with 95% humidified air and 5% CO₂.

F. Application of mechanical stretch

Cardiomyocytes were subjected to 24 hours of 10% biaxial stretch (both static and 1 Hz cyclic mode) using the setup (Fig. 20) as previously described (Kaunas et al., 2005). The central region of the membrane, which was confined with the aid
of a Teflon insert, was coated with 10 µg/ml fibronectin (Sigma, St. Louis, MO, USA) overnight and washed with sterile PBS. Cardiomyocytes were seeded on the coated region of the membrane and allowed to spread overnight. For experiments on confluent cultures, the cells were seeded at 90%-100% confluence and cultured for 2 days. Strains were measured by imaging the displacement of particles adhered to the central 4 × 4 cm surface of the membrane with a charge-coupled device (CCD) camera and a microscope (Nikon OptiPho-2) with a ×4 objective and by analyzing the particle displacements offline. Strains parallel and perpendicular to the principal stretch direction were 0.099 ± 0.008 and 0.005 ± 0.007, respectively (mean ± SD), thus resulting in essentially a 10% biaxial stretch.

Figure 20. Stretch system for static and cyclic strain. (A) The top view of the stretch chamber indicates the direction of the strain and (B) the side view shows the cells under resting (top) and stretched (bottom) states driven by the mechanical indentation.

G. Measurement of projected cellular areas

Individual video frames were digitized (Xceed Technology, Chesterfield, MI) and analyzed using image analysis software (NIH Image, v. 1.61). A discrete representation of the cell shape was obtained from digitized frames by using the mouse and the freehand selection tool to manually outline the cells. The coordinates of the cell outlines were used to compute the projected cell area.
Results

A. Identification of E-Tmod Splice Variants in NRC

Utilizing the forward primer within the rat E-Tmod exon 0 (E0) and the reverse primer downstream from the exon, PCR products (Fig. 21A) were generated and subcloned into PCR 2.1 TOPO vector (Invitrogen, Carlsbad, CA). Products generated from these primers corresponded to two E-Tmod isoforms (Fig. 21B) spanning E0-E9 (1166-bp) and E0/3-9 (850-bp). Western blot analysis confirmed the presence of protein products encoded by both E-Tmod splice variants, the 41 kDa (E0-E9) from the myofibril and the 28-kDa (E0/3-9) species in the cytosolic portion of the cardiomyocytes (Fig. 21C). The 480-bp PCR product generated in Figure 20A did not correspond to any part of the E-Tmod sequence thus was ruled out as a member of the E-Tmod splice variant.

Figure 21. (A) PCR products generated from cardiomyocytes with rat E-Tmod primers. (B) the ~1250 bp and the ~850 bp products (arrowed) were subcloned in PCR 2.1 TOPO vector was revealed and sequence verified to be distinct splice variants spanning E0-E9 and E0/3-9 regions of the E-Tmod gene. (C) Western blot of cardiomyocytes whole lysate confirmed the presence of both the 41-kD and 28-kD splice variants at the protein level.

B. Comparison of E-Tmod sequence in human, mouse and rat

The exon arrangement of the E-Tmod gene in mouse, rat and human were aligned side by side in Figure 22. Exons of rat and human were compared to the corresponding mouse sequence, with the percentage similarities indicated inside
each exon box. Genomic organizations of E-Tmod in these three organisms are essentially identical, with all of them contain nine coding exons (E1-E9) of identical size and an upstream non-coding exon (E0) more than 20 kb away. Similarities between the rodent species are higher than those between human and either mouse or rat, suggesting E-Tmod is more conserved within the rodents. As this is the first time that E-Tmod transcripts are characterized in rat, we have demonstrated the knowledge learned from the neonate are applicable to the cardiomyocytes of human and mouse as expected.

![Exon organizations of mouse, rat and human E-Tmod genes](image)

Figure 22. Exon organizations of mouse, rat and human E-Tmod genes. cDNA Sequence similarities between individual exon of rat and human were computed against the mouse sequence. The percentage identity are shown inside the exon boxes, with the length of sequence listed below the mouse exons.

C. Hypertrophic response of NRT under strain

Strain-induced hypertrophy in NRC has been a well-characterized phenomenon involving activation of SRC, EGFR and PI-3K pathways (Miller et al., 2000, Browe et al., 2003 and Sugden et al., 2003). Our experiments showed that in vitro treatment with 10% strain for 24 hours elevated the mRNA level of hypertrophic marker, atrial natriuretic peptide (ANP), in both static and cyclic stretch groups (Fig. 23A). Furthermore, projected areas of both static and cyclic stretched cardiomyocytes became noticeably larger (Fig. 23B and 23C); indicating adaptive hypertrophic response of NRC had been achieved by utilizing the strain setup described in Figure 20.
Figure 23. (A) Morphology of neonate rat cardiomyocytes in control, static (S.) and cyclic (C.) stretched groups observed under the phase-contrast microscope. (B) Averaged projected cell area (unit = µm²) of cardiomyocytes, with number of cells calculated in each group shown in the bars. (C) GAPDH-normalized expressions of cardiac hypertrophic marker, ANP, in control and two modes of stretched samples (n = 3 in each group).

D. Hypertrophic response of neonate rat cardiomyocytes under laminar shear

Ventricular cardiomyocytes are known to be exposed to fluid shear in the myocardium during cardiac contractions (Dou et al., 2003). Similar to the effect of static and cyclic strain, application of shear flow sled to an elevation of the ANP in the cardiomyocytes (Fig. 24). Under the short duration of the shear conditioning (1 hour), overall morphology of the sheared cardiomyocytes did not lead to significant change.
Figure 24. Expression pattern of hypertrophic marker, ANP in control and sheared cardiomyocytes. Asterisk denotes statistical significant change (P < 0.05).

**E. Up-regulation of E-Tmod in neonate rat cardiomyocytes by shear and static stretch**

Three different modes of mechanical treatments that mimicked distinct physiological stimulations triggered E-Tmod response in different ways. While short-term laminar shear raised both total and myofibril E-Tmod mRNA expressions (Fig. 25A), a static strain selectively increases the level of myofibril E-Tmod in neonate rat cardiomyocytes (Fig. 25B). Cyclin strain did not alter the E-Tmod level after 24 hours of treatment (Fig. 25C).

Figure 25. E-Tmod expression under (A) 1 hour of 12 dyn/cm² laminar shear, (B) 24 hour of 10% static biaxial stretch, and (C) 24 hour of 10% cyclic biaxial stretch. Expression values in all plots were measured by quantitative PCR and normalized to GAPDH level. M: E2 utilized by V1-9 + V0/3-9 transcripts; T: E6 utilized by all E-Tmod transcripts.

**F. Up-regulation of Actin and TM in neonate rat cardiomyocytes by shear and static stretch**

As both laminar shear and static strain increased the myofibril E-Tmod expressions (Fig. 25), the corresponding levels of cardiac actin (ACTC) and α-
tropomyosin (TPM-1), which constitute the main components of the thin-filament capped by E-Tmod, were also found to be significantly higher relative to the non-stimulated controls (Fig.26).

Figure 26. Cardiac actin (ACTC) and tropomyosin (TPM) expressions in sheared cardiomyocytes (A), and static-stretched cardiomyocytes (B). The levels of gene expression (Y-axis) are normalized to internal control GPADH.

G. Upregulation of E-Tmod and ANP in neonate rat cardiomyocytes by static stretch after siRNA knockdown

To investigate if decreased amount of E-Tmod affects the cardiomyocyte to undergo hypertrophy, we treated the cardiomyocytes with siRNA that targeted the E-Tmod exon 2 of the E-Tmod gene and performed static strain. The siRNA treatment knocked down an average of 36% of the E2, and 32% of the total E6 (exons utilized by all E-Tmod transcripts) in the static control as shown in Figure 27A. Stimulation of the siRNA knockdown with static strain lead to an increase in ANP expression (Fig. 27B), similar (but to a lesser degree) to the pattern observed in the non-siRNA treated cells reported in Fig. 26B.
Discussion

It has been widely accepted that 65-80% of mammalian genes can undergo alternative splicing, and as much as 50% of human genetic diseases arise from mutations related to splicing, including cancers, growth hormone deficiency and myotonic dystrophy (Garcia-Blanco et al., 2004). Hence, the study of most given gene will not be completed without the inclusion of the splice variant analysis. As we investigate the roles of E-Tmod in neonate rat cardiomyocytes, two distinct splice variants were cloned and characterized. The genomic organization and the protein localization of E-Tmod variants appeared to be identical in mouse and rat, which both contain the 1077-bp transcript encoding the 41-kDa product specifically localized at the myofibril and the 850-bp transcript encoding the 28-kDa specie primarily found in the cytosol.

Cardiomyocytes are subjected to constant hemodynamic forces in vivo and are able to undergo cytoskeletal remodeling in response to specific mechanical stimulations (Hoshijima, 2006). E-Tmod is a pointed-end capping protein of sarcomeric thin filaments that modulates actin polymerization, its role in myocyte...
responses to different modes of mechanical stress is the focus of the current study. The results from our *in vitro* experiments showed that E-Tmod expression can be altered by multiple mechanical stimulations, and that the splice variants of E-Tmod do not respond to the stimulus equally. Notably, expressions of myofibril and cytosolic E-Tmod variants rose significantly after laminar shear, suggesting that overall level of E-Tmod (both myofibril and cytosolic variants) elevated under the influence of shear. The increase in spontaneous contraction rate of the cardiomyocytes by shear agrees with the earlier report (Lorenzen-Schmidt *et al.*, 2006). Since a higher actin turnover is required to cope with the faster sarcomeric protein degradation under the more active contraction state (Russel *et al.*, 2000), the up-regulation of sarcomeric E-Tmod, in the shear-induced fast-beating of cardiomyocytes would meet this requirement.

Cyclic strain is considered to be a physiological stimulus to cardiomyocytes that mimics the *in vivo* cardiac environment (Yu and Russel, 2005). Application of biaxial cyclic-stretch to neonate rat cardiomyocytes has been shown known to induce hypertrophic response in prior reports (Yu and Russell, 2005; Sugden 2003, and Pimentel *et al.*, 2006). With 10% cyclic biaxial stretch treatment for the 24-hr duration, we were able to reproduce the hypertrophic growth of the cardiomyocytes verified by the protein to DNA ratio and elevation of hypertrophic marker ANP. Interestingly, under this 10% cyclic stretch, the level of neither myofibril nor cytosolic E-Tmod changed. Since the expression levels were not absolute but normalized to the GAPDH controls, we can only conclude that E-Tmod levels remained un-altered in relative terms and does not appeared to be specifically involved with cellular hypertrophy under cyclic stretch.
An increase in static strain is often associated with the stress-overload regions of cardiac tissues (Buerger et al., 2006). Various cardiovascular risk factors, including elevated blood pressure, narrowed coronary arteries, and pulmonary hypertension, are known to increase local strain in ventricular walls (Neyses and Pelzer, 1995; Bingisser et al., 1994). As we subjected the myocytes to the 24-hr biaxial strain, the myofibril E-Tmod level was induced noticeably together with the hypertrophic responses. Specifically, the expression of the myofibril E-Tmod was raised to a higher degree relative to the total (combined of myofibril and cytosolic) E-Tmod, indicating a preferential upregulation of this particular species. Given the fact that the two E-Tmod variants are localized in different compartments within the cardiomyocytes and do not share all of the same epitopes, it is likely that alternative variants of E-Tmod have distinct functions and therefore respond to the static-stretch differently.

Mechanical stimulation of both control and siRNA knockdown myocytes revealed that E-Tmod could be induced to higher levels by static stretch and shear. These results confirmed that the expression of E-Tmod is up-regulated during the hypertrophic process and that partial knockdown of the E-Tmod does not affect the robustness of the cardiomyocytes to undergo mechanically induced hypertrophy.

Comparison of the responses of E-Tmod under three different modes of mechanical stimulations revealed an interesting trend of expression regulation. Under the conditions that mimicked disease states of hypertrophy like (static-strain) and elevated shear of 12 dyn/cm² (relative to the 50 mdyn/cm² in vivo during systole (Schmid-Shonbein, G.W. 1999)), the myofibril E-Tmod levels increased significantly relative to the non-stimulated controls. These results indicated the modulation and adaptation of E-Tmod activities play an important roles in cardiac
remodeling process, especially under the mechanical stimuli deviated from a normal *in vivo* conditions. In contrast, the cyclic strain treatment that mimick more closely to the physiological state, did not alter the normalized E-Tmod expression.

Although both static and cyclic strains were able to induce cellular hypertrophy which required the addition of sarcomeric components, the cyclic strain may have induced the GAPDH and E-Tmod to the same proportion and hence the normalized value is comparable to the un-stimulated control group.

In both shear and static-strain groups of cardiomyocytes where myofibril E-Tmod expression were increased, the cardiac actin (ACTC) and the tropomyosin (TPM1) levels also rose in response to these mechanical stimulations. The simultaneous up-regulation of the three genes suggest that the thin-filament components (ACTC and TPM1) and the capping unit (E-Tmod) work in a synergistic manner under these specific mechanical stimulations.

Contrary to the popular hypothesis that expression level of sarcomeric genes should be up-regulated as a result of sarcomere additoin during cellular hypertrophy (Yu and Russel, 2005), we have demonstrated that mRNA levels of E-Tmod is not altered by cyclic stretch despite that hypertrophy is induced. This finding shows regulation of E-Tmod is not governed by the hypertrophy process, but rather by specific types of mechanical stimuli such as laminar shear and static stretch. According to the tensegrity model proposed by the Ingber group, a cell can adapt to pressure and loads by modulating the polymerization of actin filaments in cytoskeleton. The established role of E-Tmod in actin capping, echo with the present study that its mRNA level are altered by selective, non-physiological mechanical stimuli support the notion that E-Tmod plays a role in specific feedback and mechanosensing in cardiomyocytes.
Chapter V

Regulation of E-Tmod during cardiac myofibrillogenesis

Abstract

In the present study, we characterized the temporal expressions of the genes encoding the members of cardiac thin filament complex during myofibrillogenesis. Specifically, cardiac actin (ACTC), α-tropomyosin (TPM1), and splice variants of E-Tmod were profiled from embryonic to adult stages of the mouse heart. Our analysis of early-stage embryonic heart (day post coitus (dpc) 8 to 9.5) revealed a dramatic switch of the main E-Tmod population from cytosolic to myofibril type around the onset of the cardiac contraction. The level of the myofibril E-Tmod then rose in parallel with α-tropomyosin (α-TM) and correlated quantitatively with the increasing contraction rate and peak ventricular pressure throughout cardiogenesis. The organ culture of the dpc 9.5 E-Tmod−/− heart showed contractile activities initially but ceased completely within 72 hrs. These results indicate that E-Tmod is necessary for the maintenance of the cardiac contraction, but not for its initiation.

Introduction

Cardiac myofibrillogenesis is the process by which sarcomeric proteins are expressed and integrated to form the striated, functional structure known as myofibrils. Embryonic cardiomyocytes are unique that they can continue to divide mitotically while simultaneously synthesize myofibrillar proteins. In higher vertebrates and mammalian hearts, the population of dividing cardiomyocytes decreases as development progresses, and eventually ceases all together around the time of birth (Zak, 1974).
The basic framework of myofibril assembly consists of \( \alpha \)-actinin at the Z-disk, myomesin at the M-band with titin spanning in between. \( \alpha \)-actinin serves together with titin for the anchorage of the thin filaments, whereas myomesin helps to integrate the thick filaments with the titin filaments (Ehler \textit{et al.}, 1999; Ehler \textit{et al.}, 2004). In the early stage of cardiogenesis prior to cardiac contraction, the z-disk complex, including \( \alpha \)-actinin, F-actin and N-terminus of titin, is the first organized structure of sarcomere to take shape, followed by the maturation of the elastic (titin), the thick-filament, and eventually the thin filament system (Dabiri \textit{et al.}, 1997 and Sanger \textit{et al.}, 1986).

A key event in myofibril assembly is the regulation of actin polymerization and restriction of actin filament lengths. The uniformity of the thin filament length in myocytes is vital to the contractile function characterized by the sarcomeric length-tension curve (Huxley, 2000). The rate of fluorescence recovery after photobleaching (FRAP) of actin decreases as the myofibril matures from nonstriated to striated structure (Kishi and Shimada, 2000), suggesting that the higher exchangeability of actin monomers during the early phase of sarcomere development is due to the absence of cytoskeletal scaffolding proteins, such as capZ, Tmod, titin, and nebulette (Kishi and Shimada, 2000). The gradual formation of myofibril striation with development indicates that thin filament length-defining mechanism is developmentally regulated. Therefore, the roles of actin capping and tropomyosin binding that E-Tmod plays may contribute to this process significantly (Ito \textit{et al.}, 1995).
Material and Methods

A. Extraction of embryonic mouse hearts

Timed pregnant female mice were obtained by intercrossing C57/BL6 mice. By convention, the day a vaginal plug was seen was considered 0.5 days post-conception (E0.5), with typical gestation length being 19.5 days. Pregnant mice were euthanized by cervical dislocation. After separation of yolk sac and embryo proper whole heart was extracted from embryo proper under stereoscopic microscope (Zeiss, Stemi SV6). Animal care and experiments were done in compliance with the Animal Care Program (ACP) protocol approved by the University of California, San Diego, Animal Subjects Committee in accordance with the National Institutes of Health guidelines.

B. Culture and seeding of mouse embryonic hearts

E9 embryonic hearts were gently half-embedded in Matrigel (BD Biosciences, Franklin Lakes, NJ) of 1.5-mm thickness in the individual wells of a 96-well culture plate and gently topped by 200 ml of DMEM containing 10% FBS and 25 mM HEPES (pH 7.4). The plate was incubated in 37°C with 5% CO2. The media was switch to DMEM containing 25% rat serum with 0.05 g/L of streptomycin and fungazone, which was replaced every 12 hours.

C. Isolation of total RNA and the reverse transcription to cDNA

The cardiac tissues were homogenized immediately after harvesting and the RNA pellet was extracted in a Trizo™ (GIBCO, Paisley, UK) solution using a protocol supplied by the manufacturer. The RNA was then resuspended in nuclease-free DEPC-treated water. The quality of RNA was examined by agarose gel electrophoresis using the same amount of cardiac RNA from each animal, and
by determining the OD 260/280 ratio (all samples were > 1.8). Five µg of total RNA (1 µg/µl) from each sample was reverse transcribed at 42°C for 30 min by using the iScript cDNA synthesis kit (BioRad, Hercules, CA).

D. Isolation of total protein and separation of cellular fractions

Subcellular fractionation (membrane, cytosol, and the nuclear of the cells) was performed essentially as described previously (Nagamatsu et al., 1992) with minor modifications. The frozen hearts were minced and homogenized in 2 volumes of STE buffer (0.32 M sucrose, 10 mM Tris-HCl, pH 7.4, 1 mM EGTA, 2 mM EDTA, 5 mM NaN3, 10 mM mercaptoethanol, 20 µM leupeptin, 0.15 µM pepstatin A, 0.2 mM phenylmethanesulfonyl fluoride, 50 mM NaF, 1 mM sodium orthovanadate, 0.4 nM microcystin) in a Polytron homogenizer. The homogenates were mixed with 2 volumes of STE buffer and centrifuged (1,000 × g, 10 min) to obtain pellets. The pellet was washed once and suspended in STE buffer (nuclear fraction). The supernatant was centrifuged (100,000 × g, 60 min) to obtain the cytosolic fraction and the pellet, which was then suspended in STE buffer (membrane fraction). The nuclear fraction was solubilized in Triton buffer (1% Triton X-100, 150 mM NaCl, 10 mM Tris-HCl, pH 7.4, 1 mM EGTA, 1 mM EDTA, 0.2 mM sodium orthovanadate, 20 µM leupeptin A, 0.2 mM phenylmethanesulfonyl fluoride, 50 mM NaF, 0.4 nM microcystin LR). The fraction was centrifuged (15,000 × g, 30 min, 4 °C), and the supernatant (nuclear extract) was stored at -80 °C.

E. Protein assay and western blot

10 mg from membrane and cytosol of the cellular fractions and from the total cardiac lysate were separated by SDS-PAGE on a 10% gel (pH 8.8) and transferred onto a nitrocellulose membrane. The membrane was incubated with an anti-E-Tmod antiserum (pAb, 1:100) (Sung and Lin, 1994) or anti-E-Tmod monoclonal
antibody (G131-204 mAb, 1:150), followed by an HRP (horseradish peroxidase) conjugated goat-anti-rabbit IgG secondary antibody (1:1500) with color development or HRP conjugated goat-anti-mouse IgG (1:3000) with chemiluminescence.

F. Quantitative real-time PCR

100 ng of cDNA sample and 50 pmol of the corresponding forward and reverse primers (Qiagen, Santa Clarita, CA) were mixed with 25 µl of SYBR Green Super Mix according to the manufacturer’s protocol (BioRad, Hercules, CA), yielding a total volume of 50 µl. PCR primers (Table 8) of specific genes were designed and optimized using the MacVector program (Accelrys, San Diego, CA), and searched against the public database to confirm unique amplification products (http://www.ncbi.nlm.nih.gov). GAPDH, a housekeeping catalytic enzyme involved in glycolysis, was used as the control for the qPCR assay. Reactions in 96-well format were performed in the single color real-time PCR detection system (MyiQ, BioRad, Hercules, CA). The cycling parameters were 95°C for 15 min, followed by 40 cycles of PCR (15 sec at 94°C and 1 min at 60°C). The relative RNA expression of each sample was determined from the experimental curve normalized to the GAPDH curve and then to its corresponding wild type counterpart.
Table 8: Primers for qPCR quantifications

<table>
<thead>
<tr>
<th></th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTC</td>
<td>5’- TCCTATGTAGGTGATGAAGCCCAG -3’</td>
<td>5’- CCAGTTGGTGATAATGCGCATGTTC -3’</td>
</tr>
<tr>
<td>TPM1</td>
<td>5’- CAAAGCACATTTGCAGAAGAGGC -3’</td>
<td>5’- TGTTTCATCCATCTCTCGGCAAC -3’</td>
</tr>
<tr>
<td></td>
<td>Forward: 5’-TTCAGGAGACAAAGACACAGGCCCTC-3’</td>
<td>Reverse: 5’-TCATTTTCCAGCGTCCTCAGC-3’</td>
</tr>
<tr>
<td>E-Tmod Exon 2</td>
<td>Forward: 5’-TGGTAGGTGAAGCCTCAGGAAG-3’</td>
<td>Reverse: 5’-GAGAGTTTGGCCAAGAAGCCTCGTC-3’</td>
</tr>
<tr>
<td>E-Tmod Exon 6</td>
<td>Forward: 5’-GTCTCTCTGACACTTCAGGC-3’</td>
<td>Reverse: 5’-TCATTGTCATAACCAGAATGAGC-3’</td>
</tr>
</tbody>
</table>

G. Immunostaining

Whole mount embryonic tissues and 10-µm cryostat cardiac sections were fixed overnight in freshly prepared 4% paraformaldehyde in PBS, followed by quenching in 50 mM NH₄Cl. To detect gaps in F-actin staining in the middle of sarcomeres, and to improve signal-to-noise by reducing cytoplasmic Tmod1 staining, embryos were dissected in a relaxing buffer (Granzier et al., 1996) containing 1% Triton X-100, and incubated for 30 min on ice followed by fixation overnight in 4% paraformaldehyde in the same buffer. Embryos and sections were blocked in 4% BSA in PBS and labeled with primary antibodies overnight at 4°C, followed by fluorescent-conjugated secondary antibodies or rhodamin phalloidin. Stained tissue were analyzed on a confocal microscope (PerkinElmer UltraView Confocal Head, Hamamatsu OrcaER 1394 Camera, and Olympus IX70 Microscope). Z stacks (0.5 µm per optical section) were collected for whole mount embryos and analyzed by Image J software (National Institute of Mental Health, Bethesda, Maryland).
Results

A. Localization of E-Tmod in developing cardiac sarcomere

Immunostaining of E-Tmod counter-stained with F-actin confirmed the presence of E-Tmod during early embryonic stage (E8.5) and its localization in myofibrils at the M-line region. Analysis of the image pattern showed an increase in myofibril density (Fig. 28) and width (Fig. 29) with development, demonstrating the effect of sarcomere additions in parallel (Du et al., 2003) during myofibrillogenesis.

Figure 28. Detection and localization of E-Tmod in developing mouse hearts. The striated patterns of E-Tmod (Green) and F-actin (Red) can be visualized by E8.5. The fusion between adjacent sarcomeres can be seen in later stages of myofibril development in neonate and postnatal stages. Bars = 10 µm.

Figure 29. Myofibril width (thickness). An increase in average myofibril thickness was observed during the mouse cardiac development. Specifically, the myofibril thickness increased from 0.49 to 0.92 µm from E8.5 to the adulthood, indicating addition of sarcomeres in parallel.
B. Localization of Two E-Tmod Splice Variants in Mouse Hearts

Two splice variants of the E-Tmod gene were previously cloned in mouse tissues (Yao et al, unpublished data). PCR and western blot analysis of the mouse heart confirmed the presence of both variants (Fig 30A). Cellular fraction (Fig 30B) of the whole heart lysate indicated that the myofibril and the cytosol portions contain only the 41-kDa specie and the 28-kDa version of the E-Tmod, respectively. Since cardiac fibroblasts and endothelium lining contain no detectable amount of E-Tmod (Vera C. 2002), cardiomyocytes should be the main contributor for the E-Tmod expression observed.

Figure 30. (A) Two E-Tmod splice variants, E0-9 and E0/3-9 (B) Localization of the E-Tmod splice variants in cardiac cellular fractions. The 41-kDa product, encoded by the E0-9 spliced variant, was the only E-Tmod specie found in the myofibril fraction. The 28-kDa product, encoded by the E0-3/9 spliced variants, was the only E-Tmod located in the cytosol.

C. Temporal Expression of E-Tmod Throughout Cardiogenesis

With prior knowledge of the two alternatively spliced E-Tmod transcripts, expression patterns of exon 2 (unique to the E0-9 transcript encoding the 41-kDa myofibril variant) and exon 6 (utilized by both E0-9 and E0/3-9 variants) were measured in mouse hearts ranging from early embryonic to adult stages (Fig. 31). Temporal expression data revealed a general trend of steady but not linear increase of the level of exon 2, representing the myofibril variant, as well as the exon 6 (combined E-Tmod) throughout the course of cardiac development. It is worth
noting that the exon 2 expression increased dramatically during the early embryonic stage from E8.5 to E9.5, while the exon 6 level remained relatively constant. Expressions of both exon 2 and exon 6 remained unchanged from E9.5 to E10.5, then increased continuously beyond fetal stage until the adult phase.

Figure 31. Temporal expression of E-Tmod in mouse embryonic hearts measured by qPCR assay. E2 utilized primarily by the myofibril variant (Blue) and the combined variants (utilized by all E-Tmod transcripts + cytosol, Red) are plotted from E8.5 until adult stage. Expression level (Y-axis) represents normalized values against GAPDH.

D. Temporal Expression of TPM1 and ACTC Throughout Cardiogenesis

TPM1 encodes the sarcomeric alpha-tropomyosin molecule, which together with the cardiac actin encoded by ACTC and the pointed-end capping protein E-Tmod to form the myofibril actin thin filament complex. The expressions of TPM1 (Fig 32A) showed a similar temporal pattern of developmental increment parallel to the total E-Tmod (exon 6). The ACTC levels (Fig 32B), on the other hand, revealed a distinct expression pattern that increased more rapidly between E9.5 to E11.5, but remained at a similar level beyond that point on.
Figure 32. Temporal expressions of (A) TPM1 and (B) ACTC in mouse embryonic hearts measured by qPCR assay. The expression level (Y-axis) represents normalized value against GAPDH.

E. Correlation between temporal expressions of E-Tmod, TPM1, and ACTC

Temporal expressions of E-Tmod, TPM1, and ACTC were compared for their pairwise correlations. The side-by-side comparison showed a good correlation between E-Tmod and TPM1 (Fig. 33A and 34A), with total E-Tmod (exon 6) being slightly higher the myofibril E-Tmod (exon 2). Expression pattern of ACTC does not correlate as well with neither E-Tmod and TPM1, which is reflected by a relatively low coefficient of correlations in Figure 33B and 34B.
Figure 33 (A). Correlation between E-Tmod (exon 2) and TPM1 expressions, R = 0.97 (B). Correlation between E-Tmod (exon 2) and ACTC expressions, R = 0.74. Data points (from left to the right on both plots) were taken from E8.5, E9.5, E10.5, E11.5, E17.5, neonatal and 24-week old mouse hearts.

Figure 34. (A). Correlation between E-Tmod (exon 6) and TPM1 expressions, R = 0.99 (B). Correlation between E-Tmod (exon 6) and ACTC expressions, R = 0.81. Data points (from left to the right on both plots) were taken from E8.5, E9.5, E10.5, E11.5, E17.5, neonatal and 24-week old mouse hearts.

**F. Correlation between thin-filament genes and HR during cardiogenesis**

Cardiac contraction is the fundamental requirement for erythrogenesis and the formation of vascular network (Forouhar et al., 2006). The embryonic heart rate therefore represents a significant physiological parameter concerning developmental regulation during cardiogenesis. Several recent studies utilizing Doppler ultrasound method has established the in vivo heart rate (Fig. 35) of the mouse during embryogenesis (Nishii et al., 2006 and Leatherbury et al., 2003). The
trends of E-Tmod (Fig. 36A, 36B), TPM1 (Fig. 37A) and ACTC (Fig. 37B) expressions from E8.5 to E16.5 were plotted with the heart rate during the same period. The expression level vs. the heart rate plots (not shown) indicated that E-Tmod expression is highly correlated with the heart rate ($R = 0.99$ for exon 2 and $R = 0.98$ for exon 6), follow by the TPM1 ($R = 0.82$) and ACTC ($R = 0.76$).

Figure 35. **In vivo** heart rate of the developing mouse embryos, adapted from Nishii et al., 2006 and Leatherbury et al., 2003. Data were taken at $37^\circ$C with non-invasive Doppler ultrasound setup described in the published literature (Nishii et al., 2006).
Figure 36. Temporal heart rate and E-Tmod expressions during embryogenesis. (A) GAPDH normalized expressions of E-Tmod exon 2, representing the myofibril splice variant E0-9, are plotted with the corresponding embryonic heart rate (R = 0.99) reported by Nishii et al., 2006 and Leatherbury et al., 2003 using the Doppler ultrasound method. (B) GAPDH-normalized expressions of E-Tmod exon 6, representing the exon utilized by both E0-9 and E0/3-9 variants, are plotted with the corresponding embryonic heart rate (R=0.96). Time points collected are E8.5, E9.5, E10.5, E11.5 and E16.5 of the gestational age.
Figure 37. Temporal correlation between TPM1, ACTC expressions and the *in vivo* embryonic mouse heart rate. (A) GAPDH-normalized expressions of TPM1 at gestational age E8.5, E9.5, E10.5, E11.5 and E16.5 are plotted with the corresponding embryonic heart rate (R = 0.83) reported by Nishii *et al*., 2006 and Leatherbury *et al*., 2003 using the Doppler ultrasound method. (B) GAPDH-normalized expressions of ACTC at gestational age E8.5, E9.5, E10.5, E11.5 and E16.5 are plotted with the corresponding embryonic heart rate (R=0.76).
G. Correlation between the expressions of thin-filament genes and Fcsa

Temporal expressions of E-Tmod, TPM1 and ACTC are plotted (Fig. 38A, B, and C) with the data on Ca\(^{2+}\)-activated isometric force (Fcsa) normalized to the cross-sectional area published by Siedner using Triton-skinned myofibrils. This force generation parameter showed an excellent correlation with the myofibril E-Tmod expression (R = 0.97), followed by TPM1 (R = 0.93) but not with the ACTC expression (R = 0.29) during the gestational period from E10.5 to E16.5.

H. Correlation between temporal expressions of thin filament genes and peak ventricular pressure (VP)

Temporal expression of E-Tmod, TPM1 and ACTC are plotted (Fig. 39A, B, and C) side-by-side with the peak systolic ventricular pressure (VP) in embryonic mouse hearts determined with the micropipette method (Ishiwta et al., 2003). The correlation of VP with the expressions in the order from high to low were myofibril E-Tmod (R = 0.97), followed by TPM1 (R = 0.86) and the ACTC (R = 0.49) during the gestational period from E9.5 to E16.5.
Figure 38. Temporal expressions of (A) E-Tmod (E-2), (B) TPM1 and (C) ACTC vs. isometric force (Fcsa). Data points were taken at E10.5, E11.5, E13.5 and E16.5 of the mouse gestational stage. Data points of isometric force normalized to the cross-sectional area, Fcsa, were measured on Triton-skinned fibers by Siedner et al. 2003. Correlation of coefficients of the pairwise curves (not shown) are 0.97, 0.93 and 0.29 in (A), (B), and (C) respectively.
Figure 39. Temporal expressions of (A) E-Tmod (E), (B) TPM1 and (C) ACTC vs. peak systolic ventricular pressure. Data points were taken at E9.5, E10.5, E11.5, E13.5 and E17.5 of the gestational stage. Ventricular pressure data were reported by the Ishiwata group (2003) using servo-null micropressure system. Correlation of coefficients of the pairwise curves (not shown) are 0.97, 0.86 and 0.49 in (A), (B), and (C) respectively.
I. Expressions of $\alpha$ & $\beta$ TPM during embryonic heart development

The coiled-coil tropomyosin alpha-helix in the mature cardiac myofibrils is primarily composed of $\alpha\beta$-TM heterodimers (Bronson et al., 1982; Schachat et al., 1985). During myofibrillogenesis in mouse hearts, $\beta$-TM started as the dominating isoform. At E 17.5, $\beta$-TM was attenuated significantly, and this was accompanied by the rapid increase of the $\alpha$-TM level (Fig. 40). The proportion of $\alpha$- and $\beta$-TM then stabilized after this transition, and $\alpha$-TM remained as the primary tropomyosin isoform throughout the adult phase in cardiac myofibrils (Briggs et al., 1989).

![Figure 40. Transition of $\alpha$- and $\beta$-TM during myofibrillogenesis. Expressions of $\alpha$- (pink) and $\beta$-TM (blue) were plotted from gestational day 8.5 to 17.5. Relative expression described in Y-axis were normalized against the GAPDH level.](image)

J. Expressions of ACTC and TPM1 in $E$-$Tmod^{-/-}$ embryonic heart

ACTC and TPM1 encode the components of cardiac thin-filaments capped by E-Tmod. In the $E$-$Tmod^{-/-}$ embryonic heart, ACTC levels were reduced (Fig. 41B), and TPM1 was found to be absent (Fig. 41A), as verified by the PCR amplification (Fig. 42C). Since $E$-$Tmod^{-/-}$ embryos do not survive beyond the E10-E11 stage, cDNA has extracted from cardiac tissues of the samples collected at E9.5 of the gestational age.
Figure 41. Expressions of (A) TPM1 and (B) ACTC in wildtype and E-Tmod\(^{-/-}\) embryonic hearts. Values on y-axis are GAPDH normalized values from n = 3 samples in each group. Products of qPCR were verified in the agarose gel for correct size and specificy.

K. Expressions of cardiac troponins in E-Tmod\(^{-/-}\) embryos

Expression analysis revealed that genes (\(TnnI\), \(TnnC1\) and \(TnnT2\)) encoding components of the troponin complex in cardiac sarcomere are expressed at lower levels in \(E\text{-}Tmod^{+/-}\) compared to the \(E\text{-}Tmod^{+/+}\) embryos (Fig. 42). The troponin complex located in the cardiac actin thin filaments is known to play essential roles in \(Ca^{2+}\) dependent contractile function (Solaro and Arteaga, 2007).

L. Cardiac contraction in wildtype and E-Tmod\(^{-/-}\) cultured embryonic hearts

To investigate the role of E-Tmod in cardiac contraction, whole hearts collected from both \(E\text{-}Tmod^{+/-}\) and \(E\text{-}Tmod^{-/-}\) embryos at gestational day 9.5 were cultured and studied (Fig. 43). The organ culture of the dpc 9.5 \(E\text{-}Tmod^{-/-}\)
heart showed contractile activities initially but ceased completely within 72 hrs, in contrast, the $E-Tmod^{+/+}$ samples exhibited steady contractions beyond 7 days (Fig. 44), many even reached up to 8 weeks (data not shown). These results indicate that E-Tmod is necessary for the maintenance of the cardiac contraction, but not for its initiation.

Figure 42. Expression of cardiac troponin genes in $E-Tmod^{+/+}$ and $E-Tmod^{-/-}$ embryo propers. Expressions are measured by qPCR (n = 3) with the GAPDH normalized values shown (Y-axis). WT: $E-Tmod^{+/+}$ KO: $E-Tmod^{-/-}$.

Figure 43. Whole heart organ culture of $E-Tmod^{+/+}$ (top row) and $E-Tmod^{-/-}$ (bottom row). Hearts were extracted from E9.5 of the embryos.
Figure 44. Contraction rate of cultured embryonic hearts extracted from \textit{E-Tmod}\textsuperscript{+/–} (blue) and \textit{E-Tmod}\textsuperscript{−/−} (red) samples. Frequency on the Y-axis is indicated at beats/min.

**Discussion**

The present study characterized the expression pattern of E-Tmod during mouse cardiogenesis for the first time. The two cardiac-specific splice variants of E-Tmod, one localized primarily in the myofibrils (41 kDa) and the other in the cytosol (28 kDa), were identified both at the transcriptional and the translational levels in developing mouse hearts.

While basal levels of both E-Tmod variants can be detected at the early phase of embryonic cardiac tube (E8), expression level of the myofibril E-Tmod, as well as the total E-Tmod (myofibril + cytosolic), increased steadily throughout the course of myofibrillogenesis. Among major sarcomeric components (myosin heavy chain, cardiac \(\alpha\)-actin, and \(\alpha\)-tropomyosin), the timing in which the switching of the dominating E-Tmod species from cytosolic splice variant to the myofibril variant appeared to take place around the onset of embryonic cardiac contraction. Furthermore, temporal expression of E-Tmod, specifically the myofibril splice
variant level, correlates strongly ($r = 0.98$) with the *in vivo* heart rate (HR) observed by us and others (Nishii *et al.*, 2006 and Leatherbury *et al.*, 2003).

ACTC, TPM1 and E-Tmod are three genes that encode major components of the thin filament complex in cardiac sarcomere. The temporal expression of these genes during myofibrillogenesis differed in that TPM1 and E-Tmod correlated well with each other ($r = 0.96$), while ACTC showed no significant correlation and thus a distinct pattern by itself. The ACTC level stayed relatively constant between E8 to E11, and plateaued around E13.5 until birth. Our findings of the transcription changes are in agreements with the pattern of the protein expression of MHC and actin reported earlier (Siedner *et al.*, 2003).

As cardiac development accelerates from embryonic to fetal and postnatal stages, functional features such as heart rate, peak/end diastolic ventricular pressure (Tanaka *et al.*, 1997; Nishii and Shibata, 2006), as well as myofibril isometric force increase rapidly throughout this period (Siedner *et al.*, 2003). These functions provide the direct drive for the formation of the vascular network (Huxley, 2000), organogenesis (Ruttler *et al.*, 1973), and the embryonic circulation vital for fetal life.

By correlating the temporal gene expressions of the major thin filament components with the functional phenotypes of the developing heart, several novel observations were noted. First, the genes encoding the backbones of cardiac thin filaments (ACTC) correlate well with the myofibril density, but not with the HR nor with the peak systolic ventricular pressure. It has been reported that structural organization during myofibrillogenesis up to the multicellular alignment of myocytes appears to develop considerably more slowly than expression of MHC and actin (Siedner *et al.* 2003), and this agrees with our finding that ACTC level
increased and plateaued before the developmental increases in functional parameters of embryonic hearts.

The Ca$^{2+}$ activated isometric force normalized to the cross-sectional area (Fcsa) is an important indicator for ventricular function (Godt et al., 1991). Fcsa of the Trion-skinned fibers increases about 5 fold from E10.5 to E19.5 during the fetal stage of myofibrillogenesis (Siedner et al., 2003). The nearly linear increment of Fcsa from E11.5 to E19.5 correlates closely with E-Tmod expression (R = 0.98) as well as with TPM1 expression (R = 0.95). However, the levels of ACTC remain fairly constant between E13.5 to E19.5, indicating that an increase in force-generating ability during prenatal heart development is much less correlated to the regulation expression of the cardiac actin than E-Tmod and TPM. This observation is in agreement with previous studies reporting concentrations of both MHC and actin proteins remain relatively constant while Fcsa increases during the prenatal stage of mouse and chicken hearts (Lim et al., 1983; Godt et al., 1991 and Siedner et al., 2003).

Ventricular filling during diastole is an important determinant of cardiac performance and reflects the interaction of ventricular active relaxation, passive compliance, and filling load (Ishiwata et al., 2003). Both the developmental organizations of myofibrils and the developmentally regulated increases in the expression of calcium regulatory protein genes likely result in progressive increases in the rate of relaxation of compact myocardium. Recently, measurements of peak systolic ventricular pressure in embryonic mouse hearts using a servo-null micropressure system (Ishiwta et al., 2003) were shown to increase linearly from E9.5 to E14.5. The increment in ventricular pressure correlates well with the myofibril E-Tmod expression (R = 0.97) during the same
period of cardiogenesis. Because E-Tmod expression is highly correlated with the embryonic HR, and it has been shown that ventricular filling characteristics correlate with gestational age and HR in the human fetus (Harada et al., 1997), it is not surprising that E-Tmod expression also parallels the ventricular pressure, similar to its relation to the pattern of HR.

Three-dimensional reconstruction from serial sections of murine ventricular walls revealed a 2 fold increase in density of aligned myofibers from E12 to E16 (McLean et al., 1989), which is higher than the mere 20% increase in actomyosin (MHC and actin) concentration (Siedner et al., 2003), but comparable to the 2-fold increase of the myofibril E-Tmod expression from E11.5 to E17.5. The level of E-Tmod therefore also correlates well with myofibril density, and rises at a much faster rate than that of actomyosin accumulation during this period. This observation implicates a relatively late alignment of multicellular myocytes accompanied by the structural organization involving the E-Tmod capping activity. Our finding is in accordance with the previous report (Siedner et al., 2003) of the elevation of actin and myosin heavy chain (Myh7) prior to the sarcomeric maturation.

In summary, despite the fact that the proliferating capability of cardiomyocytes decreases temporally and eventually stops by the time the fetus is born (Rumyantsev, 1977; Olson and Schneider, 2003), E-Tmod expression increases from early embryonic through fetal and continues to the adult stage in the mouse heart. This finding suggests that the transcriptional activity of E-Tmod is developmentally and differentially regulated, instead of simply proportional to the cell proliferation which only occurs during the embryonic stage. The temporal elevations of myofibril E-Tmod splice variants correlate very well with the cardiac
parameters of Fcsa, heartrate, and peak ventricular pressure, and with the expression of another thin-filament component α-TM. This finding is significant in providing the genetic evidence for the critical role of E-Tmod in the correlation between developmental changes in ventricular diastolic function and changes in ventricular myoarchitecture (Huang et al., 2003). Expression analysis throughout myofibrillogenesis in E-Tmod deficient mice (Chapter 2) and in mechanical stress-induced hypertrophy (Chapter 4) demonstrated for the first time that alteration in E-Tmod directly affects the expressions of TM, actin, and troponin.
**Future Work**

The findings of the synergist coordination of the thin-filament components and their functions and the absence of $\alpha$-TM in the $E$-$Tmod^{-/-}$ embryos at gestational day 8.5 and 9.5 raised the possibilities that expression of $\alpha$-TM is either knocked-out or delayed as a result of the targeted disruption of the $E$-$Tmod$ gene. Although $\alpha$-TM is not a highly expressed tropomyosin species in the early developmental stage of the cardiac tissue, it has been shown (Fig. 41) to become dominating TM after gestational day 17.5 and remains highly specific in the cardiac sarcomeric thin filaments. The collective evidence that $\alpha$-TM deficiency in the $E$-$Tmod$ null mice, cardiac phenotype are identical between *in vivo* $\alpha$-$TM^{-/-}$ (Rethinasamy *et al.*, 1998) and $E$-$Tmod^{-/-}$, and that tension is required for cardiac-myofibillogenesis (Ehler and Perriard, 2000) support the possibility that $E$-$Tmod$ is required to form the thin-filament stabilizing complex with the presence of $\alpha$-TM, and *vice versa*. This actin thin-filament complex may be important for the transmission and distribution forces generated by cardiac contraction and external mechanical stimuli along the myofibrils, which in term produces a positive feedback to reinforce both $\alpha$-TM and $E$-$Tmod$ expressions. The proposed mechanism provides a logical explanation for the simultaneous up-regulation of both genes during cardiogenesis as contractile function gradually enhances and under mechanical stimulations of static strain and shear.

To further explore this novel mechanism that potentially governs the mechanoregulation of myofibillogenesis and the function of sarcomeric thin filament, we will study 1) the synergistic and compensatory relations between homologs of tropomodulins and tropomyosins; 2) whether the loss of contractile function in $E$-
Tmod-deficient embryonic cardiac progenitor cells is caused by the alternations in cell fate, retardation in growth, or apoptosis, and 3) whether mechano-sensing and force transduction along the myofibrils are necessary to constitute the feed-forward cues to up-regulate the synthesis of sarcomeric components. Further investigations in these directions will complement and extend the findings from our current study, particularly to provide a broader scheme with regard to how sarcomeric thin filament structure components make adjustments through sensing of mechano-stimuli and feedback mechanisms during myofibrillogenesis. A complete picture of sarcomeric adaptation in relation to myocardium hypertrophy and control of contraction control is needed for the understanding of cardiac myopathies and ectopic pacing irregularities, which represent some of the most common clinical problems in the field of cardiovascular medicine.
Appendix

Development of the Data Processing Algorithms and an Integrated Microarray Analysis Software Package

Objective

The aim of this section is to explore and develop the analytical techniques to improve the microarray data processing. The components of all the algorithms, including the interface as well as the mathematical modules, were written and bundled in C# (C sharp) coding language using Microsoft’s Dot-Net (.NET) platform (http://www.microsoft.com/net/). This microarray software package is available at: http://be-web.ucsd.edu/faculty/area/chien_lab/personnel.html#lian.

Introduction

Microarray technology is rapidly becoming a common tool in both academia and industry of the genomic research community. While the rate of microarray related market in U.S. has been grown at an annual compounded rate of 63% (Frost & Sullivan, 2003) with projected revenue to reach $2 billion by 2008, most of the technology providers portrayed this technology with enthusiastic and optimistic statements. Although one cannot deny microarray technology’s strength in parallel high-throughput capability and efficiency, majority of publications fail to address microarray limitations and shortcomings that could result in inappropriate applications and misinterpretation of results.

A major problematic issue associated with this technology involves interference (unspecific hybridization). The given amount of DNA probe deposited on each spot averages a mere 50ng (10,000 copies) for most platforms; with such a small amount, hybridization is greatly affected by noise from various
sources. Previous publication reported spatial variation along on a single array can range from 30% to 40% using the same set of probes (Goryachev et al., 2001), further quality control assessment experiments also suggested any single microarray output is subject to substantial variability (Lee et al., 2000).

In addition to intrinsic noise issues unique to the characteristics of the microarray system, variability of the biological assays is known to exceed the chemical assays due to the complexity of the signal and hormonal molecules. Results from characterizing normal variance in mouse gene expression (Pritchard et al., 2001) emphasize the requirement for rigorous experimental design and analysis for microarray studies involving complex tissues.

Many advocates of the microarray technology attributed its great sensitivity as one major merit. Nucleotide-based microarray platforms using fluorescent reporters are capable of detecting signals comprised of mere $10^6$ molecules or 2 pg of mRNA (Yue et al., 2001). Nevertheless, this is simply a miniaturization of the Southern blot platform and offers no advantage over the conventional detection threshold. On a per-molecule target to probe basis, microarray reactions hybridize at an efficiency equivalent to all conventional blotting methods, unless active hybridization mixing and increase of reaction surface ratio is achieved.

Determination of experiment quality based on microarray images is a particularly challenging task. To decide whether a given array has been hybridized successfully and to include selected signal spots for subsequent analysis is often subjective to experimenter’s bias. One common practice to judge the successfulness of microarray hybridization is using the abundance of hybridized signals. Commercial array manufacturer Affymetrix specifically states “the expected normal hybridization percentage of Genome arrays lies between 35
percent to 65 percent,” the rationale being not all genes at the transcriptional level can be actively synthesized at any given time point. Therefore, in contrary to the common sense, arrays with more than 2/3 of spots showing hybridized signals should be suspected of containing large numbers of false signals and artifacts. On the other hand, arrays with fewer than 40 percent hybridized genes may reflect valid expression profiles.

Procedures involved in extracting microarray data require many steps. Only through scanning, quantification, raw data treatment and normalization can meaningful, yet preliminary, expression results be obtained. Each of the processing steps requires choice of parameters, methods, and, from a statistical point of view, introduction of variations into the final tabulation. In other words, error within the already noisy microarray system is being amplified and propagated through the data-mining pipeline.

Lack of microarray analysis standard operating procedures (SOP) has been the heaviest criticism toward the microarray community. If identical array slides yield different intensity measurements when scanned at slightly different power or PMT settings and other handling procedures that can’t be quantified often produce even more problematic variations, the repeatability of gene expression results reported by array experiments is damaged.

A movement initiated by the Microarray Gene Expression Data (MGED) society to define standardized microarray procedures is currently being advocated. Adopted by major journals such as Science and Nature, experimenters submitting manuscripts containing microarray results need to comply with Minimum Information About a Microarray Experiment (MIAME). This makes the disclosure of relevant experimental and analysis procedures (approved by Microarray Gene
Expression Data (MGED) Society) as mandatory practice for submitting microarray publications.

**Development of microarray analysis pipeline**

Post-experimental procedures converting successfully hybridized microarray to annotated gene expression can be achieved in many different ways. Lacking standardized protocols and minimally optimized algorithms often leads non-repeatable results and false observations.

The framework involves in meaningfully quantify the array result required the following stages 1). Array scanning and image quantification 2). Treatment of raw data 3). Statistical processes defining criteria of genes with interests.

Despite the availability of the MIAME guidelines (Brazma *et al.*, 2001), the current microarray working group governing body, MGED, has certified and endorsed a rather limited numbers of generic algorithms. Showing an obvious unwillingness to recommend a full set of standardized microarray analysis procedure due to the varying nature of array platforms. Hence current state of microarray community relies individual core facilities, software vendors, and bioinformatic units to define specific rather than general set of procedures to satisfy the analytical phase.

The following paragraphs summarize the rationales behind the logic flow and algorithm components developed in-house specifically to optimize the microarray data analysis procedures.

**A. Image acquisition and spot quality control**

Hybridized arrays were scanned at two or more different settings in order to increase the coverage spectrum of both low and high end of signals. Scanning should be performed from lower to higher laser excitation power settings to
minimized quenching loss of fluorescent signals. Scanning equipment maker (Packard Bioscience, Model ScanArray 4000) recommends varying scanning setting by the adjustment of laser power instead of photo multiplier tube (PMT) scale unless majority of signal spots are too weak which require artificial amplification. PMT indiscriminately amplifies both feature and noise signals, which can easily overshadow low expression spots with unspecific artifact signals (e.g. background or quenching effects).

**Array signal quantified by different PMT scanning settings**

![Comparison between low and high Laser Power Scanning](image1)

![Adjusted and Extending the Saturation Curve by Normalizing with common genes](image2)

Figure A1. Incorporate high and low scanning settings by normalizing to the common genes; a non-saturated curve (right) with extended range can be obtained (Data Source: WIBE mouse cDNA self-self hybridization test slide, 11/28/02)

A successful scanning operation produces raw data covering entire distribution of detectable signal range. Intensity signals are defined by the resolution of the image (for common 16 bit scanner the intensity spectrum ranged from $2^0 = 1$ to $2^{16} = 65536$ and is unit-less therefore not absolute. The distribution of the scanned raw data often populates heavily in the lower expression regions at linear scale but otherwise symmetric (normally distributed) at log scale. If a noticeable number of genes expressed at the level of saturation then indication of inappropriate scan setting is apparent.
The first level of quality control can be exercised during the image acquisition stage. Based on physical criteria, automation tasks can assign “flag” information of individual probe according to its size, alignment, signal to noise ratio and boundary characteristics. In defining the qualify spots for subsequent analysis, the following standards were applied to distinguish the quality data spots:

(1). No pixels are greater than the background threshold.

(2). The feature diameter found is greater than the lesser of three nominal diameters (as set in the block properties) or the diameter that would cause it to overlap an adjacent feature of nominal diameter.

(3). The feature diameter goes outside the bounds proscribed in the Analysis tab of the pre-set parameter.

(4). The feature has less than 6 pixels (at 10 micron scanning resolution) above threshold.

(5). The feature has been found in a position that would overlap an adjacent feature.
B. Raw data treatment

This is a crucial step of microarray information processing, and the least transparent stage among existing protocols. To process the raw data freshly converted from an image, a sequence of actions needs to be taken. This involves background treatment and subtraction, bringing the control and treatment channels
to “comparable” level through normalization; determining the unreliable signal spots via statistical and error analysis, then making a decision either to remove or replace those values. If replicated experiments are performed, additional normalization between arrays should be performed to reduce variations contributed by chip bias.

C. Background Correction

Given backgrounds can be defined in many different ways, pixel intensities included in the background regions can be segmented by various methods; determination of background intensity from signal pixels is capable of producing valid yet non-repeatable results. Normally, background signals are magnitudes smaller than hybridized signals. Under such circumstances, expression values with high S/N ratios do not affect by background signals. However, for genes that are transcribing at low abundance, the noise level become relative significant. Background pattern needs to be profiled with high precision/fidelity in order to achieve accurate gene expression result.

At this stage it is appropriate to introduce background treatment that filters abnormally noisy signals. Such practice is considered a SOP in the field of digital signal processing (DSP). In principle, measurable background-signal is contributed by 1), Genetic noise signature resulted from ambient fluorophore, likely caused by unspecific hybridization, incomplete washing or un-hybridized residual targets on the array; 2), False inclusion of non-background pixels during segmentation process. The nature of background sources suggested the pattern of true background signals ought to be independent of hybridized probe signals. Therefore can be processed separately without being combined with the feature channels. It should be noted a correctly processed and optimized array exhibits a
rather uniform background signature across the hybridized area. Uneven spatial-distribution of noise profile suggests experimental condition requires further improvement.

Compared with Fourier high-pass filter that’s suitable to smooth artifacts of digital circuit signals, it’s reasonable from image processing point of view to apply moving average operation to the background signals reported by the raw microarray data. This procedure effectively reduces extreme values caused by false inclusion of probe signals, and presents a more uniform array background profile especially improve the validity at low expression regions. An option of assigning centroid-based weighted coefficients on a linear moving average algorithm sometimes can better reduce effects of artifact noise or local extrema.

**D. Background correction procedure**

Conventionally, pixels near the bounding box edge are taken to be background pixels, and thus the mean gray-level of these pixels provides an estimation of the local background intensity. Due to large number of subjects on each slide and the limited resolution on each individual spots, this method become inaccurate when bonding box size is close to 10 x 10 pixels, or the target fills entire bounding box. Fluorescent background is typically modeled by a Gaussian process. For example, if a larger area is chosen (e.g., a 30 x 30 box centered at a particular target), the gray-level histogram within the box is usually unimodal since the majority of the background pixel values are similar while the target pixel values spread up to very high gray levels. The location of the mode of the histogram, therefore, provides the mean local background intensity and the left tail of the histogram will provide the spread (standard deviation) of the background intensity.
E. The moving average algorithm

Among many of the useful applications, moving average is also a prototype of the finite impulse response filter, the most common type of filter used in computer-based instrumentation. In cases where a given waveform is cluttered with noise, where a mean needs to be extracted from periodic signal, or where a slowly drifting baseline needs to be eliminated from a higher frequency signal, a moving average filter may be applied to achieve the desired result. As the hybridization and washing process of microarray experiments are essentially characterized by the transport and diffusion properties of the fluids, regionalized microarray background patterns DO NOT have abrupt, sharp-cutoff edges in which local background algorithm assumes. Given that moving average algorithm accounts for both local background signal and neighboring background signal information, it minimized the error contributed by a single source of measurement, subsequently produce a more reliable estimation of true background surrounding the isolated signal spot.

The most fundamental shape of the moving average algorithm takes the shape of:

$$\sum_{i,j} B_{ni} B_{ri}$$

Which accounts for local background intensity (at i,j) and neighboring (at I ± r, j ± r) background intensity, then average these signals equally as the center of the background shifts. The weighted moving average value can be represented by:

$$\sum_{i,j} W_{ki} B_{i,j} ; \sum W_k = 1$$

This algorithm can accommodate a flexible range of radius, and carries variable weight (W) on each terms so that the user can dynamically define how the newly adjusted background value reflects the local and regional information.
F. Removal of background signals

Once the background pixels are defined and intensity profile is smoothed, either median or mean values of segmented pixels associated with each probe feature become the baseline for signal subtraction. Although the subtraction process is straightforward, dealing with features/spots with near or sub-zero net intensity require special attention. From a biological point of view a gene cannot be negatively expressed. Arithmetically, genes in the control population expressed at zero or near zero net signals can indefinitely inflate the ratio calculation, heavily skewed and falsely report up-regulation findings at erroneous magnitude. In order to reduce variations and loss of data due to marginal control signals, identify such and correction operation become a necessity to ensure quality data analysis.
G. Marginal/Low quality data handling

Once the microarray features are removed of the background signals, it is common to observe sub-zero values not suitable for normalization and ratio tabulations. While many lowly or non-expressed genes may express signals below background values or detection limits, such data category is considered with low computational value, which needs to be replaced or removed from subsequent analysis.

From signal variation point of view, another category of inferior quality signal is defined by low S/N ratios. From a error estimation point of view a measured intensity can be defined by:

$$\beta + \mu e^\eta + \epsilon$$

Where $\beta$ = feature intensity, $\mu$ = concentration of target, $\eta$ = proportional error and $\epsilon$ = baseline error

This unique representation of microarray spot signal is based on the popular Rocke-Lorenzato model, which classifies error components of any target-probe spot to either base-error ($\epsilon$) or proportional error ($\eta$). If the ratio of within spot (pixel intensity) standard deviation over mean pixel intensity exceeds a critical threshold, the data point is considered dominated by noise and should be excluded from analysis.

On the 2-color microarray system, a number of options can be exercised to reduce the effect of the low quality signals. From our experience, spots with both control and treatment signals failed to qualify with minimum S/N need to be removed from subsequent analysis. Genes with control channel below S/N should also be excluded because ratio computation of such cases often give rise to un-proportionally large expression values not only unreliable but can skew the
normalization results. The case in which treatment/experiment channel is below threshold can be rescued by replacing the values with the lowest dateable signal value or the minimum S/N of the array. Since the expression ratios of such genes are low in nature, a substitution of the numerator value does not skew the ratio computation very much.

It should be noted that removing the spot specific data could eliminate real expressions, while artificial substituting the low quality data could generate false positive signals there weren’t present in real experiment. Therefore, experimenters should be conscious about such possibilities and utilize replicated experiments to safeguard such biases.

**H. Normalization of microarray data**

Normalization is an essential and a necessary step toward microarray analysis. As the microarray gene expressions are always reported in “relative terms”, the concept of normalization is especially critical within the context of data analysis. While single channel microarray signals need to be normalized by housekeeping genes and internal controls (i.e. Affymetrix and CodeLink), multi-color arrays required normalization to the control channel in order to obtain biologically meaningful gene expression values.

The rational behind normalization is simple: matching the unit-less signal intensity with a reference point to obtain biological meaningful expression. It is important to emphasize that raw intensity reported by microarray scanner is not meaningful unless a relevant control is provided. The raw intensity can be manipulated by array surface properties, hybridization and washing conditions, not to mention scanning parameter setting in laser power and PMT level along with many other factors. The development of normalization methods has been
documented since the early days of microarray history, and has reached a mature stage in which standard protocols become widely accepted by the scientific community. Building on existed normalization platforms, we propose additional algorithms to refine conventional normalization operations, including more robust global and local-scale normalization within the same array, plus additional between-arrays normalization algorithm that further reduce slide/chip related biases.

I. Within array normalization: global scale

Standardized global normalization matches the treatment and control distribution's median or mean at log scale, and transform the normalized signals back to linear scale for subsequent analysis.

The simplest channel-wise global normalization assumes the channels being biased with a constant factor $k$ in which

$$\log(T/C) \xrightarrow{\text{Transformation}} \log(T/C) - N = \log[T/(k \cdot C)]$$

where $T = $ Treatment channel

$C = $ Control channel

$N = \log(k) = $ normalization factor = median or mean of the log intensity ratios between 2 channels

It is important to perform microarray global normalization under log scale for several reasons. Due to the nature of genome-wide expression pattern, most microarray expression signals span across a wide range of intensity spectrum, with great majority of the population resides in the very low range (Fig. A4). In order to compress the spread between extreme ends of signals, and more importantly, to
normalize the signals according to their proportional intensities, global scale normalizing operation need to be carried out under logarithmic scales.

The following log transform equations demonstrate that linear shift by a constant factor at logarithmical scale is equivalent to an exponential shift in the linear scale. This example illustrates normalization at log scale does not equivalent to normalizing the linear data by a constant.

\[ x \cdot \log_2 A \Rightarrow 2^{x \log_2 A} = 2^{\log_2 A^x} = A^x \]

\[ x \cdot \log_2 B \Rightarrow 2^{x \log_2 B} = 2^{\log_2 B^x} = B^x \]

Equations above described the log transform of microarray data by a global normalization factor \( x \), where \( A \) and \( B \) are the \(<Treatment>\) and \(<Control>\) raw signal intensities respectively.

Figure A4. Linear and log distribution of single channel microarray intensity profile: power law suggests exponential nature of microarray data. (Data Source: Cy3: \( E-Tmod^{+/−} \) cardiac tissue total RNA hybridized to Operon Oligo MouseArray, 05/27/03)
Figure A5. Global normalization between channels in log scale to match median.

Standardized global normalization matches the treatment and control distribution's median or mean in log scale, and transform the normalized signals back to linear scale for subsequent analysis. Such transformation procedure is considered superior than direct median/mean matching on the linear non-log scale, but still require fundamental assumptions of 1), Array data to exhibit near Poisson distribution pattern in log scale and; 2), Majority of genes to be highly regulated (non-differentially expressed between treatment and control groups).

Figure A6. Treatment vs. Control Plots of raw (right) and globally normalized (left) microarray data (Data Source: Operon Oligo MouseArray, 05/27/03)
Cy3: E-Tmod+/- KO whole heart total RNA
Cy5: E-Tmod+/+ Wildtype whole heart total RNA, 30μg each
J. Within array normalization: local scale

Often times, array data of different signal strength represent varying significance and confidence. Therefore, using a single, global normalization factor is insufficient to remove the spatial or pin-related variations, which are considered local properties (Quackenbush et al., 2002). The need to normalize different regions of the intensity data with more robust and non-linear mechanism is well reckoned; hence we introduce a local normalization scheme as a part of our standard data treatment procedure in the following paragraphs.

The first step to inspect for labeling bias of any microarray data is to transform the raw signals to M-A plot. M-A lot is a graphical way to see ratios and fluorescence intensity at the same time (Dudoit et al. Statistica Sinica 2002.) in which ratios are represented by A in the Y-axis and total intensities are plotted in the X-axis as M according to the following definitions:

\[ A = \frac{1}{2} (\log(C_{5}) + \log(C_{3})) \]

\[ M = \log(C_{5} / C_{3}) \]

Or from Treatment and control channel point of view:

\[ A = \frac{1}{2} (\log(T) + \log(C)) \]

\[ M = \log(T / C) \]

The merits of representing data in M-A plot format include showing eventual non-linear unwanted dependence between ratios and fluorescence intensities, and showing that using only ratios is a naive way to identify differentially expressed genes.

Representing microarray data with the ratios or log-ratio plots cannot visualize well the systematic dependence of the ratio on intensity. In order to draw biologically significant expressions rather on the strength of the hybridization
signal, the M-A plot shows if it is necessary to use a non-linear fitting for normalization.

**K. Local Regression Smoothing Procedure Using N-degree Polynomial Fitting**

1. Partition the entire span of A (total intensity axis) into small local divisions
2. For each data point, the regression weight is computed by:

   \[ \omega_i = \left(1 - \frac{|M - M_i|}{d(M_i)}\right)^3 \]

   \[ M_i = \frac{T_i}{C_i} \]

   where \( M \) is the predictor gene expression ratio associated with the response value to be smoothed, \( M_i \) are the nearest neighbors of \( M \) as defined by the span, and \( d(M) \) is the distance along the abscissa from \( M \) to the most distant predictor value within the span.
3. Within the local span window, the smoothed value \( M'_i = \omega_i \cdot M_i \)

   ![Figure A7](image)

   **Figure A7.** Multi-degree polynomial regression smoother for local normalization. The data window above illustrated a weighted linear least squares regression-fitting procedure.

   This normalization scheme has the characteristics in which data point to be smoothed has the largest weight and the most influence on the fit. In addition,
data points outside the span have zero weight and no influence on the fit. Unlike previously published LOWESS (locally Weighted Scatter Smoother) method (Chen et al., 2003) that uses a first-degree polynomial regression, a higher degree of polynomial regression is introduced to improve the continuity of the M-A profile illustrated below:

![Figure A8](image)

Figure A8. To minimize intensity related bias and spatial/print tip variations, this non-LOWESS intensity based normalization method calculates 3\textsuperscript{rd} polynomial normalization factor for each data point on the MA plot, the smoothed individual points accordingly.

Red dots: Before normalization; Blue dots: After normalization

Data Source: Operon Oligo MouseArray, 05/27/03. Cy3(G): E-Tmod\textsuperscript{+/−} whole heart total RNA, 30\textmu{}g. Cy5(R): Wildtype whole heart total RNA, 30\textmu{}g.

In contrast to the moving average smoothing process, the span is predefined by the number of partitions and never changes. For instance, if the user smoothes the data point with the smallest predictor value, the shape of the weight function is truncated by one half, the leftmost data point in the span has the largest weight, and all the neighboring points are to the right of the smoothed value.
Combining global and intensity-based local normalization the overall procedure can be summarized as:

$$\text{Log (T/C) \xrightarrow{\text{Transformation}} \log(T/C) - N(A) = \log\left[\frac{T}{(k(A) \cdot C)}\right]}$$

where $N(A) =$ Intensity dependent normalization factor = local polynomial regression fit to the MA plot.

L. Between array normalizations

Despite constant improvements on array qualities, even commercial slides can exhibit significant variability as far as the baseline and the distribution range of the signals are concerned (Piper et al., 2002). As illustrated by the replicated E-Tmod knockout experiments ($n = 6$) profile below, the spread of the individual experiments varied quite noticeably from one another even under the log scale. This prompted us to reduce inter-slide variability by introducing further normalization procedures before combing and comparing array results.

Figure A9. Comparison of ratio distributions of 6 replicated array experiments. Data source: E-Tmod$^+/-$/E-Tmod$^+/-$ cardiac expression: 12/12/02-12/12/03
Variability between arrays has long been documented but remains little exploited relative to other aspects of microarray error assessment due to 1) prohibitive cost of large replicated sets required to carry out meaningful normalization; 2) expression ratios from individual arrays are considered independent as well as biologically “absolute”; 3) lack of a consensus standard and point of reference.

To overcome these challenges, we first recognize that the intensity distribution of any high-density array represents a unique, independent profile of the experiment. Consequently, ratio expressions derived from Treatment/Control signals is qualitatively (referring to the regulation trend) significant in a biological sense, and should remain consistent after the inter-slide normalization procedure.

However, within the context of not changing the regulation trend, the quantitative expression ratios of any given gene may be dependent to the non-experimental related factors such as S/N ratio, detection and labeling sensitivities as well as other factors. Therefore, it is legitimate to adjust and normalize the value of the expression ratios of replicated arrays by matching the ratios between slides to minimize such variability.

The introduction of “Outlier Removed Range-Matching (ORRM)” normalization is an effort to remove the scaled-related variations between replicated arrays. To carry out this normalization procedure, a simple step-by-step guideline is defined to be the following:
1. Take log for all the ratio columns involved

2. Determine the 95% percentile (2.5% to 97.5%) range of ratios of the array to be normalized, quantified as R1

3. Determine the 95% percentile range of the ratios averaged from all other replicates, quantified as R2

4. R2/R1 is the normalization factor

5. Multiple R2/R1 to log ratios of array population to be normalized

6. Convert normalized log ratios back to non-log scale

   Essentially, the ORRM method matches the outlier-excluded ratio range of individual array to the average outlier-excluded ratio range compiled from all other replicated arrays. Because normalization is carried out under log scale, this procedure does not alter the symmetry of Ratio = 1 (under log scale = 0) position, therefore regulation trend (either up or down) remains the same for all genes. In addition, given R2 does not include contribution form the range of the array being normalized to, this procedure does not introduce bias from R1 nor is a biasing attempt to minimize the variations by normalizing to a component of itself.

---

Figure A10. ORRM method matches the scales of expression ratio (log-ranged) to the mean distributions of other replicated. This procedure reduces variations caused by arrays with uneven sensitivity range and minimizes the extrema/outrlier data effects.
To further validate this novel algorithm, a published set of microarray data previously analyzed by the Stanford biostatistics group (Speed et al., 2002) was tested. According to the following data summary, the ORRM method effectively reduces the statistical variations between replicates and improves pair-wise correlations between experiments.

Table A1. Reduction of SD and CV with ORRM treatment

<table>
<thead>
<tr>
<th>Original Data</th>
<th>Range-Matched Norm.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>sd</td>
</tr>
<tr>
<td>0.515227</td>
<td>0.191225</td>
</tr>
<tr>
<td>0.554839</td>
<td>0.173968</td>
</tr>
<tr>
<td>0.595035</td>
<td>0.171736</td>
</tr>
<tr>
<td>0.563705</td>
<td>0.094282</td>
</tr>
<tr>
<td>0.579376</td>
<td>0.141518</td>
</tr>
<tr>
<td>0.573442</td>
<td>0.059934</td>
</tr>
<tr>
<td>0.579658</td>
<td>0.134371</td>
</tr>
<tr>
<td>0.581422</td>
<td>0.175221</td>
</tr>
<tr>
<td>0.583161</td>
<td>0.219641</td>
</tr>
<tr>
<td>0.583249</td>
<td>0.046776</td>
</tr>
<tr>
<td>0.583509</td>
<td>0.043948</td>
</tr>
<tr>
<td>0.587633</td>
<td>0.129726</td>
</tr>
</tbody>
</table>

Values in which either P-Value or CV is Larger
Values in which either P-Value or CV is Smaller

Selected down-regulated genes before and using ORRM cross slide normalization method, where CV and t-test profile are improved with a general trend of reduction in variations.

Table A2. Reduction of p-value with ORRM treatment

<table>
<thead>
<tr>
<th>Original Data</th>
<th>Range-Matched Norm.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>sd</td>
</tr>
<tr>
<td>2.030209</td>
<td>1.450654</td>
</tr>
<tr>
<td>2.034659</td>
<td>0.529576</td>
</tr>
<tr>
<td>2.051316</td>
<td>1.622559</td>
</tr>
<tr>
<td>2.091044</td>
<td>0.829688</td>
</tr>
<tr>
<td>2.096886</td>
<td>0.585857</td>
</tr>
<tr>
<td>2.119613</td>
<td>0.780887</td>
</tr>
<tr>
<td>2.134862</td>
<td>2.030157</td>
</tr>
<tr>
<td>2.147957</td>
<td>0.401341</td>
</tr>
<tr>
<td>2.16081</td>
<td>1.667722</td>
</tr>
<tr>
<td>2.181026</td>
<td>1.152070</td>
</tr>
<tr>
<td>2.264367</td>
<td>2.658559</td>
</tr>
<tr>
<td>2.474693</td>
<td>2.716509</td>
</tr>
</tbody>
</table>

Values in which either P-Value < 0.05
Values in which either P-Value < 0.05 and improved
Values in which either P-Value improved but still > 0.05

Selected up-regulated genes without (Original Data) and using ORRM normalization treatment, where CV and t-test profile are improved with a general trend of reduction.
Table A3. Improved correlation between arrays with ORRM treatment

<table>
<thead>
<tr>
<th></th>
<th>$R^2(\text{before})$</th>
<th>$R^2(\text{after})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Array 1</td>
<td>0.403881</td>
<td>0.446899</td>
</tr>
<tr>
<td>Array 2</td>
<td>0.50447</td>
<td>0.51451</td>
</tr>
<tr>
<td>Array 3</td>
<td>0.620901</td>
<td>0.621197</td>
</tr>
<tr>
<td>Array 4</td>
<td>0.50878</td>
<td>0.525709</td>
</tr>
<tr>
<td>Array 5</td>
<td>0.610857</td>
<td>0.632018</td>
</tr>
<tr>
<td>Array 6</td>
<td>0.590279</td>
<td>0.61487</td>
</tr>
<tr>
<td>Array 7</td>
<td>0.507642</td>
<td>0.534041</td>
</tr>
</tbody>
</table>

Average pairwise correlation coefficient of array combinations. Pairwise correlation between individual array vs. averaged array expression ratio profile. The correlation coefficient is improved in all slides after applying ORRM normalization method.

Data source: M.J.Callow & E.M Rubin of Stanford.
CY3: Experiment channel: Apolipoprotein AI (apo AI) KO mouse.
CY5: Control C57B1/6 wildtype mouse.

M. Data Mining and Filter

After background reduction, normalization, and outlier removal (via quality control steps) procedures, information generated from microarray data is considered properly treated ready to probed for biological significance. At this stage, although the individual set of experimental data is quality checked, noise-removed and normalized between channels (and between experiments), it’s still difficult for the experimenters to interpret the data without further tabulation and mining.

In order to rank the data according to its biological significance and statistical confidence levels, computation of mean ratio followed by standard deviation (SD), coefficient of variation (CV) as well as t-test P value from all replicated sets are straightforward steps to perform.

In additional to the conventional statistical methods to characterize data variations, a novel score describing the agreement between replicated data is
introduced. By comparing individual experiment’s gene expression ratios to the averaged ratios, this qualitative score compiles the accumulative information correlating the regulation trend (up or down) between all replicated experiments. For instance, if 3 out of 4 replicated samples show the upward regulation pattern agreeing with the averaged ratio (also an up-regulation), an “agreement score” is defined as 3 (number of experiments showing same regulation patterned with the mean ratio) over 4 (total number of experiments) or simply 75% as illustrated in the example below. If the sample size among the replicated population is large, a reasonable cutoff would be (n-2)/n representing a logical tolerance of at least 1 outlier value on each end of the extremes.

Table A4. Example of agreement score computation

<table>
<thead>
<tr>
<th>Mean Ratio</th>
<th>Exp 1</th>
<th>Exp 2</th>
<th>Exp 3</th>
<th>Exp 4</th>
<th>Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>1.2</td>
<td>1.8</td>
<td>2.5</td>
<td>0.5</td>
<td>¾ = 75%</td>
</tr>
</tbody>
</table>

The need to introduce the “agreement score” is to safeguard the loose statistical criteria (such as CV = 50%) of which much microarray publications have been solely relied on (McCormick et al., 2001). In addition, often times the T-test P value is too sensitive to tolerate a single outlier value, consequently making false negative calls on otherwise valuable data points. Therefore, a non-parametric, logical filter such as the “agreement score” on top to the combination of statistical filtration criteria with adjustable parameters is capable of making the microarray data mining more intelligent.

The following table summarizes our recommending criteria and the range of parameters effective in prioritize the replicated microarray data according to its statistical confidence. This portion of work enables experimenters to sort out the most repeatable gene expression groups from the entire dataset.
Table A5. Choice of filter criteria and cutoff parameters.

<table>
<thead>
<tr>
<th>Test</th>
<th>Criteria</th>
<th>Cutoff Range</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Coefficient of Variation (CV)</td>
<td>50%</td>
<td>Parametric</td>
</tr>
<tr>
<td>B</td>
<td>T-test</td>
<td>0.05 to 0.1 (sample size dependent)</td>
<td>Parametric</td>
</tr>
<tr>
<td>C</td>
<td>Agreement Score</td>
<td>50%-100%</td>
<td>Non-Parametric</td>
</tr>
<tr>
<td>D</td>
<td>Flag Score</td>
<td>Above Marginal</td>
<td>Non-Parametric</td>
</tr>
</tbody>
</table>

Once the data set is filtered according to its quality measures, defining biological significance is the final and the most important step toward the microarray analysis task. Depending on the purpose of the microarray experiment design, experimenters’ definition of biological/experimental significance can vary tremendously. For instance, screening for the active genes under specific experimental conditions, the most up and down-regulated genes become the center of focus. For the purpose of profiling sample specific expression pattern or to identify functional roles of the unknown genes, cluster by expression pattern would produce more fruitful results.

In order to define the differentially regulated genes from replicated microarray data, the choice of method and cutoff criteria is required. Most publications to date employed an expression ratio cutoff of 1.5 or 2 fold induction as the minimum level of significant regulation change (Kaminski and Friedman, 2002). Needless to say such arbitrary criteria are subjective and non-scientific. A better choice than a numerical cutoff is to use the tails of the distribution curves of the entire expression population. While any array constituting large number of genes tends to exhibit an expression profile similar to the Gaussian distribution (Goryachev et al., 2001), the tail portions on the distribution curve usually well represent the extremes of the expression population in an objective manner. Since
gene expression ratio and induction change is a relative measurement, on an either highly or hardly differentially distributed array, a non-conditional ratio cutoff easily yields too many or too little genes to work with.

A more intelligent way to define differentially regulated genes would tolerate the medium of the expression ratio distribution. Since medium of the population represents the most abundant and the least-changing (no fold induction) gene pools, therefore going 1 to 2 standard deviation away from the medium point to both stretches of the distribution will effectively locate the most active (significantly regulated) genes under the curve, as illustrated below.

Figure A11. Definition of differentially expressed genes according to ratio distribution pattern.

Figure A12. Venn diagram summarizing logics of microarray filters. Combining the statistical filters and the significance criteria discussed above, a systemic Venn diagram can be generated to partition the data such as the following:
• Group 1: Region A+F+I+E: denotes replicated expression ratios showing similar regulation trend and safeguarded by the good quantitative consistency.

• Group 2: Region B+G+I+F: denotes genes with expression ratios of smallest variation and highest statistical confidence.

• Group 3: Region C+H+I+G: denotes characteristically differentially expressed genes with statistical confidence.

• Group 4: Region D+E+I+H: denotes genes with significant expression changes agreed among replicates, but the variations between numerical values may be large.

• Region I+F: Union includes all areas shared between group 1 and group 2.

• Region I+G: Union includes all areas shared between group 2 and group 3.

• Region I+H: Union includes all areas shared between group 3 and group 4.

• Region I+E: Union includes all areas shared between group 4 and group 1.

• Region F denotes the union between Group 1 and Group 2, which share the logical characteristics of both gene clusters.

• Region G denotes the union between Group 2 and Group 3, which share the logical characteristics of both gene clusters.

• Region H denotes the union between Group 3 and Group 4, which share the logical characteristics of both gene clusters.

• Region E denotes the union between Group 1 and Group 4, which share the logical characteristics of both gene clusters.

• Region I denote the group of genes meeting requirements described by all statistical, numerical and logical filters of choice of cutoff parameters.
To examine the independence of each statistical filter, close comparison between standard student t-test and CV is needed. From mathematical point of view, CV is simply the standard deviation of the replicated ratio normalized to the mean ratio. Various forms of t-test (Student, Bonferroni, Mann-Whitney…etc) such as the ones represented here:

\[
\text{t value} = \frac{1}{\sqrt{n}} \left( \frac{x - \bar{x}}{SD} \right) \quad \text{(Student t-test)}
\]

\[
\text{t value} = \frac{X_1 - X_2}{\sqrt{2s^2/\bar{n}/n}} \quad \text{(Bonferroni t-test)}
\]

Mathematical relations indicate that t value is inversely proportional to the CV \((s_{\text{wit}} \approx SD)\) but cannot be considered comparable for several reasons. Given the t value is dependent on the sample size and measured statistical variation between replicates, it is not an ideal score for values with mean ratio close to 1 \((\bar{x} \equiv 1)\). As a large portion of gene expression data tend to remain unchanged after normalization, genes with expression ratio between treatment and control closes to unity will yield t values near zero (consequently yield a very large p-values much greater than the common 0.05 cutoff). Despite expression measurement of such genes may be highly accurate and repeatable, they are often overlooked and excluded from subsequent analysis due to failure of passing the t-test. Since not all microarray experimentalist intend to look only at the most differentially expressed genes, hybridized probes exhibiting small expression change are still valuable for clustering profiling and functional identification purposes.

While the most actively-expressed genes (highly deviated from expression ratio of 1) should be screened for statistical variations through t-test, majority of other genes with smaller (but not necessary insignificant) fold-inductions are often too
sensitive to the t-test therefore should not be excluded solely by this criteria. An alternative approach to validate these genes is through the quantitative CV filter in addition to the non-parametric “agreement score” discussed in the previous paragraphs.

When CV measures the agreement between replicates by normalizing SD to the mean expression ratio, this score does not penalize samples with small number of replicates as much as the “agreement score” directly comparing the regulation trend among all replicates. In the case where most replicates express a treatment over control ratio around 1 yet the replicate size is small (n < 4), using CV 50% as cutoff criteria tend to retain many genes that would otherwise excluded by t-test (P = 0.05 cutoff) and agreement score using cutoff of (n-2)/n.

Table A6. Example of expression ratio close to 1 gene exhibiting large P value.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Exp 1 Ratio</th>
<th>Exp 2 Ratio</th>
<th>Exp 3 Ratio</th>
<th>Exp 4 Ratio</th>
<th>Exp 5 Ratio</th>
<th>Mean Ratio</th>
<th>CV</th>
<th>t-test value</th>
<th>t-test P value</th>
<th>Agreement Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ankyrin3, epithelial</td>
<td>1.07</td>
<td>1.47</td>
<td>0.67</td>
<td>0.74</td>
<td>0.97</td>
<td>0.99</td>
<td>0.32</td>
<td>0.007</td>
<td>0.81</td>
<td>60%</td>
</tr>
</tbody>
</table>

Data source: E-Tmod<sup>+-/+<sup> cardiac expression: 12/12/02-12/12/03.

**N. Knowledge-based clustering with GeneSet Enrichment Analysis (GSEA)**

Conventional microarray clustering method profiles individual gene based on expression pattern alone. Such technique relies on a single parameter and often group functionally un-related genes together. An alternative to this approach is GeneSet Enrichment Analysis (GSEA), which cluster differentially expressed, functionally related gene groups in a knowledge-based manner. Based on the idea that alterations in gene expression might manifest at the level of biological pathways or co-regulated groups of genes, rather than individual genes
(Subramanian et al., 2005), the GSEA utilizes the following algorithm to determine if a specific subset of genes (under the same functional category, for example) is differentially expressed in a co-regulated manner:

1. Define $n_a$ genes in group A, $n_b$ genes in group B.
2. Order the genes with respect to expression values.
3. Create vector $vv$ of $(n_a + n_b)$ components with value $-n_b$ at each position where a value from group A is sitting and with value $n_a$ at each position where a value from group B is sitting.
4. Calculate $yy = \text{cumsum}(vv)$.
5. Draw a starting at $(0,0)$ through points $(I,yy[I])$. The line will end in $(n_a + n_b,0)$ because $(-n_b)n_a + n_a n_b = 0$.
6. Look at $Mvv = \max \{\lmin(yy), \lmax(yy)\}$ which will be large in case of good separation between both groups.
7. Permute the vector $vv$ to get $vv^*$, calculate $yy^*$ and $Mvv^*$. Use permutation to calculate the distribution of $Mvv$ under the Null hypothesis, determine the permutation based p-value: $p = \# \{Mvv^* \geq Mvv\}/\# \text{of permutations}$.

While the subset of gene group tested by GSEA can be arbitrarily defined, the functional classification assigned by Gene Ontology (GO) consortium is one appropriate parameter to group the genes with. Since GSEA determines if a given gene group is differentially expressed by its relative rank with respect to the entire population, this method does not depend on the absolute expression of individual gene nor an arbitrary reference point (centeroid) to initiate. Clustering result in Chapter II demonstrated that GSEA algorithm is effective in generating clusters by both biological significance and expression pattern simultaneously. Furthermore, GSEA allows comparison of expression data from different experiments and even with non-identical formats, making it a robust and flexible technique in analyzing microarray-oriented datasets.
Conclusion

The nature of the highly sensitive and dense microarray gene expression platform requires dedicated methodology to extrapolate the reliable information from. The work in this chapter streamlined a statistical-based microarray analysis package based on specialized and tested algorithms.

Like any biochemical assays, measurements of microarray expression signals comprised of estimation and variability components. It is important that the expression data is being processed with proper consideration in statistical confidence, experimental and hybridizing related bias as well as logistics of data mining criteria.

The software package developed in this project addressed these issues by 1), applying a background signal treatment that improved sensitivity and accuracy of the signals especially in the low expression regions comprising majority of the data population; 2), utilizing multi-step normalization procedures to minimize labeling, intensity, and slide specific biases; 3), prioritizing and filtering data according to its statistical confidence, degree of logical agreement and biological significance.

It has been a great concern for many microarray experimentalists about the repeatability of not only the replicated microarray assays but also the results processed through different methodologies and software. Within the scope of our development and testing experience, minor to moderate numerical variations resulted from the choice of different data treatment schemes and parameters are often expected and observed. Nevertheless, biological significance and filtered results remain the same for the majority of differentially expressed genes, notable their regulation pattern and statistical confidence levels.
By developing and implementing the improved microarray analysis package, the goal of this project is not to achieve a non-supervised, complete automation pipeline. Instead, the refined algorithms and additional options accompanied with it requires further user understanding about the nature of experimental design relative to the added data mining options. Given the complexity of modern genomic research and the constant evolution of microarray technology, it is essential to invest scientific resource on replications and updated analytical methods in order to draw reliable inferences from such studies.

The installation files of this microarray analysis package can be accessed at: http://be-web.ucsd.edu/faculty/area/chien_lab/personnel.html#lian
REFERENCES


Junqueira LCU, Bignolas G, Brentani RR. (1979) “Picrosirius staining plus polarization microscopy, a specific method for collagen detection in tissue sections”. Histochem J. 11, 447-455


Kuhlman PA, Fowler VM. (1997) “Purification and characterization of an alpha 1 beta 2 isoform of CapZ from human erythrocytes: cytosolic location and inability to bind to Mg2+ ghosts suggest that erythrocyte actin filaments are capped by adducin.” Biochemistry. 4;36(44):13461-72.


