Title
Epigenetic Regulators of Cardiac Hypertrophy and Failure

Permalink
https://escholarship.org/uc/item/4tk8m559

Author
Jiang, Shanxi

Publication Date
2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Epigenetic Regulators of Cardiac Hypertrophy and Failure

A dissertation submitted in partial satisfaction of the

requirements for the degree of Doctor of Philosophy in

Molecular, Cellular and Integrative Physiology

by

Shanxi Jiang

2018

ABSTRACT OF THE DISSERTATION

Epigenetic Regulators of Cardiac Hypertrophy and Failure

by

Shanxi Jiang

Doctor of Philosophy in Molecular, Cellular and Integrative Physiology

University of California, Los Angeles, 2018

Professor Thomas M. Vondriska, Chair

Under pathological stress, an otherwise healthy heart may enter hypertrophy, a partially-reversible, compromised state wherein heart function is relatively normal although the muscle cells increase in size. Should the stress continue, however, the heart will succumb to the irreversible condition of heart failure, resulting in an inability to efficaciously pump enough blood to support bodily demands. When the heart enters states of either hypertrophy or failure, noticeable changes in chromatin accessibility and gene expression arise. Chromatin accessibility can be defined by a binary chromatin state model: heterochromatin is tightly packed and contains silenced genes, while the relatively loose conformation of euchromatin is more conductive to active gene transcription. Alternations in gene expression or epigenetic regulation are revealed using high-throughput sequencing techniques, which have been developed and rigorously applied over the last two decades. How the principles revealed from studies

of chromatin impact gene expression levels in the diseased heart is unknown. My dissertation studied the role of multiple chromatin regulator factors, as studied by high-throughput sequencing techniques, contribute to function of the normal and diseased heart. Those factors include a histone modifying enzyme, a nucleosome remodeling protein, circular RNAs and DNA methylation.

The first two chapters of my dissertation detail the functions of two chromatin remodelers identified by quantitative proteomics using mouse hypertrophy and heart failure models: Smyd1, a histone methyltransferase coding gene containing the SET and MYND domains, and Napl14, nucleosome assembly protein 1-like 4. Chapter 1 reports that the chromatin-binding protein Smyd1 restricts adult mammalian heart growth. Mice with induced knockdown of cardiac-specific Smyd1 displayed cardiomyocyte growth, organ remodeling, and declined heart function. Chapter 2 describes a possible mechanism by which histone chaperone Nap1l4 may regulate cardiac transcription in hypertrophy. As revealed by siRNA knock down, the lack of Nap1l4-mediated transcription reduces the size of neonatal rat ventricular myocytes (NRVMs) and inhibits fetal gene reprogramming induced by phenylephrine (PHE). However, when Nap1l4 is overexpressed, there is an increase in the size of NRVMs.

The latter two chapters of the dissertation describe the epigenomic changes revealed by high-throughput sequencing that could potentially affect gene expression during cardiovascular diseases. Chapter 3 explores our utilization of Ribo-Zero RNA sequencing to discover circular RNAs (circRNAs) in the heart using mouse models. We confirmed the existence of cardiac-related circRNAs including circMyocd, circRyr2, and circTtn. With the successful knockdown of circMyocd in NRVMs, we observed increased

expression of linear Myocd, indicating the circRNAs may regulate transcription of its linear counterparts. Chapter 4 characterizes DNA methylation alterations in patients undergoing coronary artery bypass grafting (CABG) using reduced representation bisulfite sequencing (RRBS) with respect to post-operative atrial fibrillation (POAF). When comparing pre-operative and post-operative epigenomic states, we found that the hypervariable CpG sites are mostly enriched in or around genes pertaining to the immune system, cellular adhesion and the cardiovascular system. Specifically, altered CpG methylation in genes coding for transforming growth factor-beta 1 (TGF-β1) may be a marker for POAF as well as pre-operative and post-operative epigenomic states. My dissertation revealed that epigenetic changes including chromatin remodelers, DNA methylation and circRNAs could affect the gene expression during heart diseases. The work will undoubtedly benefit the whole community and shed light on the translational medicine for heart failure patients.

The dissertation of Shanxi Jiang is approved.

Arjun Deb

Jingyi Li

Xinshu Xiao

Thomas M. Vondriska, Committee Chair

University of California, Los Angeles

2018

DEDICATION

This dissertation is dedicated to my parents

who always encouraged and supported me

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to first express my deepest gratitude to Dr. Thomas Vondriska for his continuous support and guidance as my mentor since I was an undergraduate exchange student in his lab in 2011. He provided the best scientific training and made me a better person. He educated and inspired me to think critically and creatively, to pay attention to details, to communicate with different people, to present my work confidently and to write grants effectively. I am extremely lucky to have worked with him and I really enjoyed my time in his lab. At the same time, I genuinely appreciated the help from all my lab members: Dr. Elaheh Karbassi, Dr. Manuel Rosa Garrido, Dr. Emma Monte, Mr. Doug Chapski, Mr. Max Cabaj, Dr. Matt Fischer, Dr. Sarah Franklin, Dr. Haodong Chen, Dr. Michelle Parvatiyar, Mr. Todd Kimball, Dr. Andrea Matlock, Dr. Scherise Mitchell-Jordan, Ms. Rachel Lopez, Ms. Elizabeth Soehalim, Ms. Irene Shih, and Ms. Celina Nishioka. In particular, I would like to thank my deskmate Elaheh for her help in almost everything, Manuel for his help in designing projects, Emma for her help in proteomics and grant applications, Max and Matt for their help in the DNA methylation project, Haodong and Doug for their help in bioinformatics, Sarah for her guidance in the Smyd1 project and Michelle for her help in the NAP1L4 project.

I would like to thank my dissertation committee for their assistance and advice: Dr. Arjun Deb, Dr. Jingyi (Jessica) Li and Dr. Xinshu (Grace) Xiao. They were always there to help me and I am really grateful for their feedback. Meanwhile, I wish to thank my collaborators for their help: Dr. Yibin Wang, Dr. Shuxun (Vincent) Ren and Dr. Matteo Pellegrini.

I also want to thank my parents for their support and my friends who gave me ideas for refining my projects at various stages of the Ph.D.

Last but not least, I would like to acknowledge the permissions from American Physiology Society and Springer Publishing to use previously published manuscripts in my dissertation. I am also grateful for the National Institutes of Health, the American Heart Association, and the Department of Anesthesiology at UCLA for their generous support of the Vondriska lab. I am thankful to be the recipient of the Chinese Council Scholarship.

<div align="center">Biographical Sketch</div>

EDUCATION

INSTITUTION & LOCATION	DEGREE	DATES ATTENDED	FIELD OF STUDY
Fudan University, Shanghai, China	Bachelor of Medicine	09/2008~07/2013	Preventive Medicine
University of California, Los Angeles	Doctor of Philosophy Student	09/2013~	Molecular Physiology

ACADEMIC AND PROFESSIONAL HONORS

2013	Chinese Scholarship Council Scholarship
2015	MCIP Student Support Award
2017	MCIP Graduate Division Block Grant

PEER REVIEWED PUBLICATIONS

Research Papers:

- **Jiang, S**, Vondriska, TM. (Agnetti et al. Editors). Sensing and remembering cellular states through chromatin. Manual of Cardiac Proteomics. (2016).

- Franklin S, Kimball T, Rasmussen T, Rosa Garrido M, Chen H, Tran T, Miller M, Gray R, **Jiang S**, Ren S, Wang Y, Tucker H, Vondriska TM. The chromatin binding protein Smyd1 restricts adult mammalian heart growth. American Journal of Physiology - Heart and Circulatory Physiology. 2016;311:H1234-47.

Abstracts:

- **Jiang S**, Karbassi E, Franklin S, Vondriska TM. The Histone Chaperone NAP1L4 Regulates Cardiac Transcription in Hypertrophy. Annual MCIP Retreat, Hollywood Roosevelt Hotel in Hollywood, 2014

- **Jiang S**, Chapski DJ, Karbassi E, Kimball T, Lopez R, Monte E, Rosa Garrido M, Shih TT, Vondriska TM. Systems Epigenomics of Cardiovascular Diseases. Department of Medicine Research Day, UCLA, 2015

- **Jiang S**, Cabaj M, Kimball T, Fischer MA, Mahajan A, Pellegrini M, Vondriska TM. Genome-wide DNA Methylation Analysis of Cardiac Surgery Patients. Annual MCIP Retreat, UCLA, 2018

INVITED LECTURES AND SEMINARS

- New Roles for Chromatin Structural Proteins and Circular RNAs in Heart Disease. Annual MCIP Retreat, UCLA, 2017

Introduction: Sensing and Remembering Cellular States through Chromatin

Shanxi Jiang and Thomas M. Vondriska

Abstract Chromatin is the means by which the same genome encodes multiple cells: it enables orderly development, normal physiology and, when it goes haywire, malfunctioning chromatin is a hallmark of disease. In the cardiovascular system, the epigenomic features controlling gene expression have recently become the focus of intense research. This chapter examines the principles of chromatin structure, details their regulation and identifies areas of rapid development in our understanding of how the genome is packaged. Also explored are the recent observations indicating that deranged epigenomic features on a genome-wide scale may underpin various cardiovascular diseases.

Keywords Heart • Vasculature • Epigenetics • Epigenomics • Genomics • Proteomics • Transcription

Introduction

Around two billion years ago, natural selection for a strategy of information storage that utilized RNA and protein to package DNA presaged the evolution of multicellularity. To get the same DNA substrate to produce different outcomes, the method or conditions of extracting information has to change. Chromatin accomplishes this task in all plants and animals: comprised of histone proteins and DNA, chromatin is the structural form of the genome in vivo, compacting the enormous chromosome molecules for storage in the nucleus. That plants and animals can achieve the spectacular range of appearance and function observed in the natural world is the result of the ability to produce highly specialized cells and organs. This specialization is achieved through chromatin and its regulation, which enables the same genomes to orchestrate diverse phenotypes throughout development and into adulthood.

The fundamental unit of chromatin is the nucleosome, consisting of two copies of four histone proteins (named H2A, H2B, H3 and H4) and entwined by 147 base pairs of DNA [1]. Interesting relationships have emerged relating primary DNA sequence to the binding of nucleosomes across a genome: however, it is now recognized that a diverse range of regulatory mechanisms control where nucleosomes reside, how they combine with each other to form higher order structures and the resulting accessibility (or lack thereof) for transcriptional machinery to interact with and express a gene. Broadly construed, DNA can be either accessible for transcription, or euchromatic, or inaccessible for transcription, heterochromatic [2]. As depicted in Fig 15.1, the landscape of chromatin features combine to facilitate or prevent gene transcription.

It is now commonly accepted that chromatin patterns are cell and developmental stage specific, underlying transcriptome changes that enable phenotype specification. Individual nucleosomes can be modified by swapping the histone variants that comprise them, by post-translational modification of those histones, by ATP-dependent processes that reorganize groups of nucleosomes in response to environmental stress or developmental cue and by alterations to the DNA, cytosine methylation in particular. Furthermore, RNA and protein can combine with nucleosomes to form higher order chromatin structures, compacting or relaxing sub-chromosomal regions. With the development of the next generation sequencing techniques, there has been an explosion in the analysis of how different proteins bind to the genome and the combinatorial patterns of protein binding that specify chromatin structure and thereby transcriptional behavior at individual loci [3]. A major challenge for biology and medicine is now to decode the logic of chromatin regulation, to enable targeting of gene expression programs for therapeutic modulation during disease as well as to use epigenomic information for improved patient stratification and diagnoses.

This chapter reviews the current understanding of chromatin regulation in disease with a particular emphasis on the cardiovascular system, describing how genetics and environment are integrated in the epigenome, ultimately controlling disease susceptibility and progression.

Basics of Chromatin Structure

Histone Variants

Although they all contain two copies each of four histone family proteins, nucleosomes are not homogeneous protein complexes. Mammalian genomes harbor dozens of histone variant genes that, when expressed, can combine to specify a range of chromatin features. In the heart, proteomic mass spectrometry has been used to quantify histone variants [4], investigating how these variants change with disease [5]. Much of what is known about the role of individual histone variants comes from loss of function studies in animal models.

Replacing histone H3, centromere protein A (CENP-A) epigenetically defines centromere organization through a process that involves RNA interference [6]. In addition to CENP-A, there are several more centromeric histone associated proteins, including CENP-B, CENP-C and CENP-T, that maintain the function and structure of CENP-A and promote centromere formation during cell division [7]. In the setting of some cancers, overexpression of CENP-A results in enrichment at non-centromeric sites, thereby altering chromatin structure as part of the disease pathogenesis [8].

Another histone H3 variant is H3.3, which is highly conserved, with only four amino acid differences from the canonical H3 in eukaryotes [9]. Localization of histone H3.3 corresponds to transcriptionally active chromatin regions with the highest turnover rate at RNA polymerase II (RNAP II) binding sites and transcription start and end sites, indicating its function in transcription initiation and termination [10]. Furthermore, the presence of H3.3 regulates repressive histone marks (e.g. H3K27me3) [11] and histone variants (e.g. H2A.Z) [12]. *De novo* synthesized H3.3 replaces the canonical histone H3 and remodels donor nuclear chromatin for gene activation during oocyte reprogramming and knockdown of H3.3 disrupts the

reprogramming process [13]. Complete deletion of H3.3 embryonic lethal due in part to p53 activation [14].

The N-terminal MacroH2A is 64 % homologous to the canonical histone H2A and contains a carboxyl-terminal ~30 kDa macrodomain [15]. MacroH2A had long been thought to regulate gene silencing, however recent chromatin immunoprecipitation (ChIP) plus DNA sequencing findings indicate that macroH2A localizes to areas of both repressive and active chromatin, having its effects on transcription by inhibiting activator binding sites (to cause repression) or inhibiting repressor sites (to cause activation) [16].

Histone variant H2AX, constituting between 2 and 25 % of total H2A protein by gel image quantification in mammals [17], has been functionally implicated in mitotic/meiotic division, stem cell development and aging [18]. Recently, it has been reported that H2AX deposition can be serve as an epigenetic mark for quality control of induced pluripotent stem cells [19]. Phosphorylation of H2AX serves as a marker of DNA damage.

H2A.B is unstable at the protein level and distant in terms of sequence homology compared with other histone H2A variants, with ~40–50 % sequence identity [20, 21]. H2A.B-containing nucleosomes wrap only 116–130 bp of DNA (rather than the conventional 147 bp) and as a result transiently associate with the genome during processes of DNA replication and repair [21]. Genome-wide analysis showed that H2A.B correlates with DNA methylation (in some scenarios a repressive marker itself) and facilitates methylation related to transcription elongation, suggesting a positive role in regulating gene expression [22].

Histone H2A.Z is a highly conserved H2A variant with about 60 % homology to canonical histone H2A, demonstrating both active and repressive transcriptional regulation. Recent findings show that H2A.Z associated with gene coding regions as well as 3' and 5' ends of genes, facilitates cryptic, antisense transcription and RNAP II regulated transcription [23, 24, 25]. H2A.Z works with Nanog, a key transcription factor for stem cell identity, in regulating pluripotency/reprogramming and serves as a biomarker for asymmetrically self-renewing cells [26, 27]. Increased protein levels of H2A.X cause cardiac hypertrophy, whereas knockdown of H2A.Z prevents pathologic cardiomyocyte growth [28].

Excluded from the nucleosome core, the last family of histone variants, known as linker histones, plays an intriguing and incompletely understood role in controlling how nucleosomes interact with each other. The lysine-rich linker histone (also referred as H1 and H5) contains a highly conserved central globular domain, a short amino-terminal tail and a long carboxyl-terminal domain, which are vital to its role in higher-order chromatin structure [29]. The globular domain of chicken linker histone H5 and Drosophila linker histone H1 have been shown to possess off-dyad and on-dyad binding respectively, perhaps due to differences in key amino acids between the proteins [30]. Linker histones can interact with both core histones and other proteins. Through mass spectrometry and microscopy, Histone H1 has been found to interact with H2A C-terminus and proteins that are involved in rDNA chromatin structure, rRNA processing and mRNA splicing [31, 32, 33]. Furthermore, knockout of all three H1 isoforms in mouse embryonic stem (ES) cells revealed roles for the protein family in gene silencing and nucleosome organization [33]. The main modification of linker

histones is phosphorylation by cyclin-dependent kinases (CDK), which is lowest in G1 and highest in G2 and Mitosis [34]. Infrared spectroscopy of phosphorylated linker histone with CDK2 showed the induction of β-structure that may result in chromatin condensation [35]. Impaired linker histones dynamics can trigger multiple diseases including cancer [36].

Notable to the heart, triple knockout of H1cH1dH1e of E9.5 showed pericardial expansion [37]. At the mRNA level, these triple knockout mice did not show alterations in cardiac specific transcripts such as Nkx2.5 and alpha myosin heavy chain [38]. In an adult model of cardiac hypertrophy, global epigenome remodeling involves changes in histone stoichiometry at the protein level: following pressure overload, the ratio of linker to core histones was decreased as compared to the healthy heart, suggesting a more permissive transcriptional environment. This interpretation is supported by genome-wide transcriptome changes in cardiac hypertrophy, which would necessitate relaxed chromatin at multiple loci. Lastly, this change was associated with a global shift of histone post-translational modifications favoring euchromatin over heterochromatin [5].

High mobility group proteins, originally shown to be structural components of chromatin in the 1970s [39] have more recently been observed to contribute to cardiac gene expression in a locus-specific manner, facilitating higher order chromatin structure [5]. These proteins may act similar to linker histones, facilitating higher order structure of chromatin by binding bent DNA and enabling compaction of tracts of nucleosomes.

Chromatin Domains

While there is widespread agreement on the importance of the nucleosome as functional unit of chromatin and its modulation by chromatin remodelers (discussed in detail below), the principles of genome organization beyond this scale are decidedly more nebulous. Let's estimate there are perhaps 20 million nucleosomes in a given nucleus and around 3 billion base pairs of DNA: how this molecular morass is organized for reproducible, timely access and repackaging is of key importance, and indeed has been the focus of intense investigation. Early observations of in vitro reconstituted DNA and histones revealed the formation of the titular '30 nm fiber', in which the linker histone teams up with several nucleosomes to form intermediate domains of packaging. More recently, 10 nm substructures have been identified, again comprised of nucleosomes plus linker histones, and suggested to be the functional units of both euchromatin and heterochromatin [40]. The folding of this 10 nm nucleosome fiber has been demonstrated to be irregular and gathered in heterogeneous groups, leading to variable chromatin structure, and challenging the long held view about a higher order 30 nm fiber secondary structure [41]. It has been found that chromatin secondary structure is affected by many factors, such as linker histones, length of linker DNA and thus spacing of nucleosomes, histone variant and histone/DNA modifications [42, 43]. In this active area of investigation, novel techniques [42] are continually being brought to bear on the question of whether there is a finite intermediate structure of chromatin, larger than a nucleosome and smaller than the chromosome.

For decades it has been recognized that the genome segments into non-random chromosome territories which may play a role (although the direction of a causal arrow is unclear in this relationship) in transcriptional programs [44]. Recent advances in next generation sequencing have expanded and textured this model in interesting ways. First, innovation in chromosome capture technologies have enabled multiple studies [45] into the endogenous structure of the mammalian genome, revealing folding principles, providing higher resolution to aforementioned chromosomal territories and establishing the fractal, self-repeating structure adopted by the cell's most complex multi-molecular complex. As further resolution has been achieved with these studies, topologically associated domains (TADs) have been described, accounting for ~90 % of genomic structure in mouse ES cells genome and averaging 880 kb in size [46]. The boundary regions of TADs are enriched for CTCF, housekeeping genes and short interspersed elements, but not histone modifications such as enhancer-related H3K4me1 and heterochromatin-related H3K9me3 [46]. TAD boundaries tend to be conserved between cell types, however within these boundaries, the histone modifications and chromatin structural proteins that decorate chromatin impose upon it transcriptional phenotypes, including all variations of active and silenced chromatin [47]. Thus TADs appear to be a structure within which epigenomic modifiers specify the transcriptome. Likewise, TADs correspond to eukaryotic replication-timing reprogram, translocating from nuclear interior during DNA synthesis (active transcription) to nuclear periphery during later replication (repressive transcription) [48]. Deletion or disruption of TADs by CRISPR/Cas genome editing leads to altered gene expression and *de novo* enhancer-promoter interactions [49].

9

Open questions in chromatin structure include: To what extent is our ability to discern intermediate chromatin features limited by the fixation and sequencing protocols currently in vogue? If TADs are shared between cell types, what is the role of global chromatin structure in cell type specific transcriptomes and phenotypes? What is the impact of genetic variation, which has been all but ignored in epigenomic studies to date, on chromatin structure?

Remodeling Chromatin for Development and Disease

To accommodate multiple cellular transcriptomes, chromatin accessibility has evolved to be highly dynamic throughout normal organismal development. This plasticity is also exploited in disease to enact abnormal transcriptional programs. Recent advances in understanding how epigenomic memories are created, remembered, erased—and, to extend the metaphor—in some cases hallucinated, has advanced our understanding of the basic biology of gene expression and shifted our understanding of disease to include aberrant chromatin structure and function.

Histone Tail Modification

While the bulk of the histone protein mass is ordered within the core nucleosome particle, each of the four variants wears a capricious amino terminus that, resisting fixed structure, is exposed to the nuclear milieu, solicitous of molecular interaction and post-translational modification. The ability of enzymes to modify histone tails to influence transcription in vivo has been known for nearly 40 years [50, 51], but recent advances in mass spectrometry has exploded the list of modifications

documented on histone tails to include virtually every known type of post-translational modification, numbering in some cases over 100 modifications on nucleosomes from a single cell type and with most but not all occurring on the soluble tails [52]. Methylation (active or repressive transcription) and acetylation (active transcription) represent two major classes and coordinate with each other [53]. Other histone modifications include phosphorylation, ubiquitination, SUMOylation and crotonylation. Through computational methods, the relationship between histone modifications and gene expression has been explored [54].

Some of these principles for histone modification-dependent regulation of gene expression and phenotype have been tested in the heart. Methylation and acetylation of histones with their associated histone modifiers (such as acetyltransferase, histone deacetylases and histone methyltransferases) have been particularly well explored [55, 56]. Other modifications, such as phosphorylation of histone H3 is involved in cardiac hypertrophy through transcriptional elongation [57]. Moreover, hyper-acetylation has been shown to change gene expression through RNA alternative slicing, thus affecting cardiac cell growth [58]. Stimulation of SUMOylation may exert a protective function on heart [59]. Genome wide analyses have identified cooperative functions of active (H3K9ac, H3K27ac, H3K4me3, and H3K79me2) and repressive (H3K9me2, H3K9me3, and H3K27me3) histone marks [60]. Using data from the ENCODE Project, Roadmap Epigenomics and several other studies, researchers found that E11.5 active enhancers can be accurately predicted by three dimensional analysis of genome-wide H3K27ac and H3K4me1 across developmental time, between tissues within an organism and for the corresponding tissue within species [61]. An algorithm named

histoneHMM (hidden markov model) has been used to predict genomic regions that affect cardiac hypertrophy, mostly focused on H3K27me3 [62]. Moreover, histone methylation levels in the heart have been found to be regulated mostly in trans, prominently for H3K4me3 [63].

ATP-Dependent Remodelers

Where nucleosomes reside along the genome affects gene expression and is thus a highly regulated process. One of the most direct mechanisms to influence nucleosome positioning in an active, stimulus responsive manner is through ATP-dependent chromatin remodeling enzymes, which are usually subdivided into four groups: SWI/SNF (switching defective/sucrose non-fermenting), ISWI (imitation switch), INO80 (inositol requiring 80) and CHD (chromodoman, helicase, DNA binding). These groups are identified by a specific, highly conserved ATPase that belongs to the SF2 helicase superfamily and can alter histone-DNA interactions through a process that consumes ATP. These proteins influence transcription as well as chromatin structure by nucleosome translocation [64], nucleosome (whole nucleosome/H2A-H2B dimers) eviction [65] and histone variant exchange (e.g. H2A.Z) [66].

In the heart, most studies have focused on SWI/SNF in the setting of development and hypertrophy, yet all four groups have been explored in some manner [55]. Chromodomain-helicase-DNA-binding protein 7 (CHD7; from the CHD family) and its mutation has been associated with CHARGE syndrome, revealing its novel function in calcium excitation-contraction coupling [67]. Probably the most well studied

family is that regulated by Brg1, a member of the BAF complex, which has been shown to be critical for fetal gene activation (myosin heavy chain switching in particular) in the mouse heart following stress [68].

DNA Methylation

Not all things chromatin occur on proteins. The first, and perhaps only truly *epigenetic* (as defined by transgenerational heritability of acquired features) mark, DNA methylation is defined as the addition of a methyl group (CH_3) to the C5 position of cytosine and usually occurs within the major groove of DNA at CpG dinucleotides The methyl group is transferred by DNA methyltransferase (DNMT) family with DNMT1 functioning as the maintenance, and DNMT3a/b as the *de novo*, DNA methyltransferase [69]. DNA methylation mostly happens in the promoter region, gene body and less in intergenic regions. Across the genome, DNA methylation and CpG density follow a bimodal distribution, with high methylation level in CpG-poor regions (CpG depletion) and low methylation level in CpG-rich (typically entails > 50 % GC) region that are commonly defined as CpG islands (CGIs) [70]. CGIs are located mainly in promoter regions and their methylation can initiate vigorous, long-term transcription repression such as X-inactivation [1].

Although DNA methylation usually associated with repressive transcription, studies show that DNA methylation, especially in gene bodies, is altered during transcriptional elongation [71]. Another long held view—that DNA methylation is highly stable—is also being reconsidered in light of the finding of active DNA demethylation (higher in gene bodies) through base excision repair proteins such as Ten-eleven

translocation (TET) [72]. Interesting, DNA methylation has been regarded as a regulator of alternative splicing possibly by influencing chromatin structure, thus affecting RNAP II recruitment and binding of heterochromatin protein 1 [73].

In the heart, DNA methylation influences cardiomyopathy and heart development. In congenital heart diseases, hypermethylation of MSX1 and GATA4 has been found [74]. In Tetralogy of Fallot, certain genes such as EGFR and TBX5 have shown significant differences in methylation status compared to unaffected individuals [75]. In patients with dilated cardiomyopathy, aberrant DNA methylation has been found in lymphocyte antigen 75 and adenosine receptor A2A, which has been further confirmed in zebrafish [76]. Also, three angiogenesis-related genes (AMOTL2, ARHGAP24 and PECAM1) have been identified that exhibited altered methylation status [77]. During development, only a small fraction of genes showed aberrant methylation: this subset, however, were highly related to cardiac specific processes when comparing developmental day E11.5 to E14.5 [78]. When comparing adult to developing mouse heart, DNA methylation is increased in active enhancers [79]. When comparing developing, mature and diseased cardiomyocytes, DNA methylation is quite dynamic, supporting a role for this modification in promoting, or responding to, condition-specific gene expression [80]. Cardiomyocytes treated with endothelin-1 (a hypertrophic agonist) showed increased DNA methylation; conversely, inhibition of DNA methylation rescued the norepinephrine-induced hypertrophy [81, 82]. Therefore, alteration of DNA methylation has been proposed as a novel therapeutic target in the heart, although the mechanisms of action, and principal targets, remain to be determined.

RNA-Based Mechanisms

Widespread application of next generation RNA sequencing technologies have dramatically increased the portion of the genome that is understood to be transcribed in a given cell type, now estimated at ~75 %. The beguiling nature of this observation is that most of these newly identified transcripts do not appear to be messenger RNAs: instead, they belong to a class of RNAs called, perhaps misleadingly, non-coding RNAs (ncRNAs) which may code small peptides/proteins, may function independently as RNA scaffolds for various cellular processes or may in fact be, as was previously thought, transcriptional noise. ncRNAs can be divided into small RNA (<200 nucleotides) and long non coding RNA (lncRNA, >200 nucleotides) [83]. Small RNAs are thought to modulate heterochromatin and gene silencing with the help of Argonaut [84]. They may also participate in alternative splicing and transcription together with other epigenetic regulation such as DNA methylation and histone modifications [85]. lncRNAs connect chromatin loci with chromatin remodelers, transcription factors and other RNAs both in cis and trans [86, 87] by binding modules nestled in their secondary structure, leading to changes of chromatin structure and nuclear organization [88]. Besides their close relationship to silenced chromatin, different classes of lncRNAs alternatively contribute to active transcription, especially transcriptional enhancer element RNA (eRNA) that binds the Mediator complex [89]. Natural antisense transcripts (NAT), another class of lncRNA, are read from the opposite strand of the mRNAs that they regulate in a complementary, cis manner.

Alternations in lncRNA levels have been implicated in cancer, skeletal defects, embryogenesis abnormalities and brain defects [90].

Roles for lncRNAs in heart development, differentiation, and disease have recently emerged. Depletion of the lncRNA *Braveheart* revealed its role in cardiovascular lineage commitment by activating MesP1 (a master transcription factor expressed in multipotent cardiac progenitor) and interacting with SUZ12 (a component of polycomb repressive complex 2(PRC2)) [91]. Targeted homozygous deletion of *Fendrr* in mouse, another lncRNA, resulted in embryonic lethality and PRC2 reduction, leading to decreased H3K27me3 and increased H3K4me3 in the promoter regions of target genes [92]. Restoration of repressed myosin heavy-chain-associated lncRNA transcripts (abbreviated as *Myheart*) in the setting of pressure overload hypertrophy protects the heart from cardiomyopathy by interacting with and inhibiting Brg1 [93]. Myocardial infarction associated transcript (*Miat*), a potential risk factor of myocardial infarction, was discovered through a case–control association study of single nucleotide polymorphism markers [94]. The relationship of microRNAs to heart and vascular diseases, has been extensively reviewed [95, 96].

circRNA, a novel class of ncRNA, is formed by backsplicing and features a covalently joined loop structure without free 3′ and 5′ ends [97]. Circular Antisense Noncoding RNA in the INK4 Locus (ANRIL) is the first cardiovascular related circRNA that correlated with INK4/ARF expression and atherosclerosis risk [98]. Using a statistical method, circular Sodium/Calcium Exchanger (NCX1) was found to increase more rapidly during fetal heart development than its linear version [99]. Recently, using human umbilical vein endothelial cells, researchers showed that circZNF292 is

16

regulated by hypoxia and displays proangiogenic activity [100], although this family of RNAs remains largely unexplored in the cardiovascular system.

New Epigenomic Techniques

Sequencing-Based Techniques

For histone variants, restriction endonuclease digestion of chromatin coupled to deep sequencing (RED-seq) offers an unbiased and sensitive method to study chromatin accessibility in nucleosome depleted regions, within nucleosome arrays and between different histone variants [101]. RED-seq is performed on permeabilized cells using restriction endonuclease digestion, DNA extraction with unbiased sonication, two separate steps of ligation of linkers, PCR amplification and sequencing. For tertiary structures such as TADs, Hi-C, that studies in situ DNA-DNA contacts, has been applied [46]. Sub-TAD structure, however, has been explored with the help of higher resolution 5C [102]. Apart from DNA-DNA interactions in shaping the three-dimensional arrangement of chromatin, the function of specific proteins, especially transcription factors, can be identified through chromatin interaction analysis with paired end tag (ChIA-PET) that involves immunoprecipitation [103]. ChIA-PET can investigate specific proteins that modulate genome organization through formaldehyde crosslinking, DNA sonication, ChIP enrichment, followed by ligation and sequencing.

For examining the occupancy profile of nucleosomes along chromatin, multiple techniques have emerged. For open chromatin, transposase-accessible chromatin using sequencing (ATAC-seq) features a simple and sensitive two-step protocol to explore the nucleosome landscape [104]. Without fixation, permeabilized gDNA is

ligated with the help of Tn5 transposase. After purification, open chromatin is amplified and sequenced. Another novel approach named methidiumpropyl-EDTA sequencing (MPE-seq) provides a sensitive method for detection the upstream open chromatin region of active promoters' transcription start sites, which can be combined with micrococcal nuclease-sequencing (MNase-seq) to generate a detailed readout of chromatin structure and regulation [105]. MPE-seq uses ferrous iron to generate DNA breakage followed by sequencing with minimal bias. For histone modifications and transcription factor binding sites, ChIP followed by qPCR or sequencing has been extensively used to map the proteins or histone modifications bound to DNA with cell fixation, DNA sonication and sequencing [106]. Reduced representation bisulfite sequencing (RRBS) is a commonly used method for DNA methylation (cytosine methylation and hydroxylmethylation) analysis. RRBS uses bisulfite to convert unmodified cytosines to uracil, thereby revealing methylation distribution. A new technique named methyl-sequencing has been developed that has the ability to detect the 5-formylcytosine that is resistant to conversion in RRBS, combined with comprehensive methylation detection level by using cytosine-methylated universal adapters [107]. For detecting the hydroxylmethylation alone, Tet-assisted bidulfite sequencing (TAB-seq) [108] and oxidative bisulfite sequencing (oxBS-seq) [109] can be employed. TAB-seq uses Tet to convert methylcytosine to carbomethylcytosine, whereas oxBS-seq applies oxidization to convert hydroxymethylcytosine to formylcytosine. For transcriptome and alternative splicing, RNA sequencing remains to be the top choice. For the relationship between lncRNA and chromatin, chromatin isolation by RNA purification (ChIRP) sequencing can be used [110]. ChIRP maps the

genomic localization of a known ncRNA by antibody pull down. Besides RNAs, nascent transcripts that bind to RNAP II can be detected through methods such as nascent transcript sequencing (NET-seq) [111].

Through whole genome sequencing, specific genes that may affect the congenital heart diseases, dilated cardiomyopathy, arrhythmia and other heart-related diseases have been detected [112]. For changes in coding sequences, exome sequencing has emerged as a powerful choice due to lower cost and simpler interpretation in contrast with whole genome sequencing. Many heart diseases, such as congenital heart diseases [113], myocardial infarction [114], and coronary heart disease [115] have been explored with these approaches. Similar to RNA sequencing, DNA methylation sequencing has been compared with normal and diseased conditions, revealing aberrant methylation in genes associated with dilated cardiomyopathy [76], cardiac fibrosis [116], and congenital heart diseases [78]. Numerous proteins have been shown, by ChIP-seq experiments, to alter their association with the genome during heart disease, development and regeneration [93, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126].

Microscopy

Using microscopy to visualize chromatin structure has tantalizing appeal, but limitations abound with regard to resolution. Light microscopy has a resolution around 200–300 nm and employs florescent probes to visualize living cells. Certain techniques, such as fluorescence recovery after photobleaching (FRAP) and fluorescence in situ hybridization, can be coupled with light microscopy. For super-

resolution microscopy, besides the antibody-based methods which have been applied to examine chromatin structure in cardiac myocytes during disease [127], new technique such as TALENs [128] and CRISPR/Cas9 [129] will enable the endogenous gene labeling for both proteins and DNA sequences that will give more reliable results that can be tailored to a specific genetic locus. Other readouts of genomic organization including RNA polymerase II, histone modifications and nascent RNA have been investigated [130, 131]. Imaging of the core histone H2B has also been used to reveal the spatial organization of chromatin fiber [132]. Cryogenic electron microscopy, with around 10 nm resolution, has been used to study chromatin structure, revealing the formation of higher-order chromatin features influenced by the presence of H1 [42].

In the heart, super-resolution microscopy, especially its application in T-tubule and calcium signaling, has been reviewed [133]. Label-free microscopy, generating light signals based on molecules' photo-physiology, has been applied to reveal cardiac-vascular interactions [134]. Intravital microscopy, which is applied to living tissues, has been used to investigate the beating heart with single cell resolution during cardiac ischemia [135] and monocyte circulation during myocardial infarction [136]. Atomic force microscopy, a type of scanning probe microscopy with resolution ranging from 0.1 to 100 μm, allows 3D structural assessment in living cells. This method has been used to detect sarcomere lengthening [137] and valve leaflet stiffness [138]. Different kinds of electron microscopy have enabled the identification of myocyte morphology [139], ultrastructure of the intercalated disc [140], morphology of mitochondria [141] and the interaction of telocytes and myocytes [142]. Frontiers for

microscopic analysis of cardiac chromatin will be breached by novel labeling reagents as well as by the application of new imaging techniques, such as 3D super resolution and light sheet microscopy.

Concluding Remarks and Perspective

Chromatin is the substrate of cellular memory—it is the way cells, and therefore organs, remember what they are. The last decade of research in genomics, epigenomics and transcriptomics have revolutionized our understanding of the mechanisms through which these different tiers of biological information interact. Despite this progress in the basic science realm, this holistic approach to biology is challenging for translation: progress requires that we come to utilize cellular networks in similar terms as the EKG…to interpret a readout from 'omics measurements akin to how chest auscultation is employed as an integral part of clinical decision making. This advance requires novel approaches to analyzing big data and, most importantly, investigations of epigenomic regulation and epigenomic susceptibility directly in human populations.

Notes

Acknowledgements

Heart Association, Thermo Fisher Scientific and the Department of Anesthesiology in the David Geffen School of Medicine at UCLA.
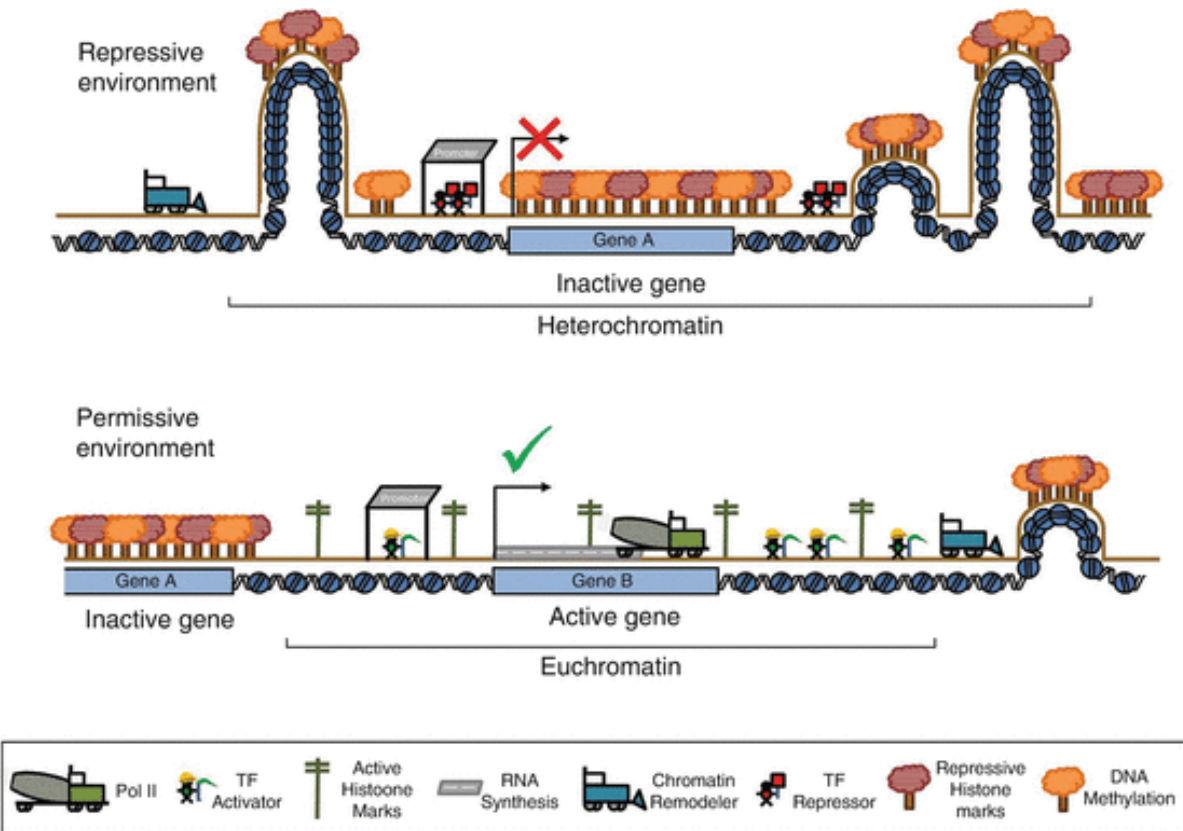
Fig 15.1 Chromatin landscaping. At individual loci, a balance of features promoting or inhibiting chromatin accessibility determines whether transcription occurs. As described in the text, these features decorate individual nucleosomes and combine to influence gene expression and chromatin structure on the scale of genes, chromatin territories and the entire genome

References

1. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet. 2003;33(Suppl):245–54.

2. Kouzarides T. Chromatin modifications and their function. Cell. 2007;128:693–705.

3. Soon WW, Hariharan M, Snyder MP. High-throughput sequencing for biology and medicine. Mol Syst Biol. 2013;9:640.

4. Franklin S, Zhang MJ, Chen H, Paulsson AK, Mitchell-Jordan SA, Li Y, Ping P, Vondriska TM. Specialized compartments of cardiac nuclei exhibit distinct proteomic anatomy. Mol Cell Proteomics MCP. 2011;10:M110.000703.

5. Franklin S, Chen H, Mitchell-Jordan S, Ren S, Wang Y, Vondriska TM. Quantitative analysis of the chromatin proteome in disease reveals remodeling principles and identifies high mobility group protein b2 as a regulator of hypertrophic growth. Mol Cell Proteomics MCP. 2012;11:M111.014258.

6. Black BE, Cleveland DW. Epigenetic centromere propagation and the nature of cenp-a nucleosomes. Cell. 2011;144:471–9.

7. Logsdon GA, Barrey EJ, Bassett EA, DeNizio JE, Guo LY, Panchenko T, Dawicki-McKenna JM, Heun P, Black BE. Both tails and the centromere targeting domain of cenp-a are required for centromere establishment. J Cell Biol. 2015;208:521–31.

8. Lacoste N, Woolfe A, Tachiwana H, Garea AV, Barth T, Cantaloube S, Kurumizaka H, Imhof A, Almouzni G. Mislocalization of the centromeric histone variant cenh3/cenp-a in human cells depends on the chaperone daxx. Mol Cell. 2014;53:631–44.

9. Elsaesser SJ, Goldberg AD, Allis CD. New functions for an old variant: no substitute for histone h3.3. Curr Opin Genet Dev. 2010;20:110–7.

10. Ha M, Kraushaar DC, Zhao K. Genome-wide analysis of h3.3 dissociation reveals high nucleosome turnover at distal regulatory regions of embryonic stem cells. Epigenetics Chromatin. 2014;7:38.

11. Weth O, Paprotka C, Gunther K, Schulte A, Baierl M, Leers J, Galjart N, Renkawitz R. Ctcf induces histone variant incorporation, erases the h3k27me3 histone mark and opens chromatin. Nucleic Acids Res. 2014;42:11941–51.

12. Chen P, Wang Y, Li G. Dynamics of histone variant h3.3 and its coregulation with h2a.Z at enhancers and promoters. Nucleus. 2014;5:21–7.

13. Wen D, Banaszynski LA, Liu Y, Geng F, Noh KM, Xiang J, Elemento O, Rosenwaks Z, Allis CD, Rafii S. Histone variant h3.3 is an essential maternal factor for oocyte reprogramming. Proc Natl Acad Sci U S A. 2014;111:7325–30.

14. Jang CW, Shibata Y, Starmer J, Yee D, Magnuson T. Histone h3.3 maintains genome integrity during mammalian development. Genes Dev. 2015;29:1377–92.

15. Pehrson JR, Fuji RN. Evolutionary conservation of histone macroh2a subtypes and domains. Nucleic Acids Res. 1998;26:2837–42.

16. Lavigne MD, Vatsellas G, Polyzos A, Mantouvalou E, Sianidis G, Maraziotis I, Agelopoulos M, Thanos D. Composite macroh2a/nrf-1 nucleosomes suppress noise and generate robustness in gene expression. Cell Rep. 2015;11:1090–101.

17. Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM. DNA double-stranded breaks induce histone h2ax phosphorylation on serine 139. J Biol Chem. 1998;273:5858–68.

18. Turinetto V, Giachino C. Multiple facets of histone variant h2ax: a DNA double-strand-break marker with several biological functions. Nucleic Acids Res. 2015;43:2489–98.

19. Wu T, Liu Y, Wen D, Tseng Z, Tahmasian M, Zhong M, Rafii S, Stadtfeld M, Hochedlinger K, Xiao A. Histone variant h2a.X deposition pattern serves as a functional epigenetic mark for distinguishing the developmental potentials of ipscs. Cell Stem Cell. 2014;15:281–94.

20. Chadwick BP, Willard HF. A novel chromatin protein, distantly related to histone h2a, is largely excluded from the inactive x chromosome. J Cell Biol. 2001;152:375–84.

21. Arimura Y, Kimura H, Oda T, Sato K, Osakabe A, Tachiwana H, Sato Y, Kinugasa Y, Ikura T, Sugiyama M, Sato M, Kurumizaka H. Structural basis of a nucleosome containing histone h2a.B/h2a.Bbd that transiently associates with reorganized chromatin. Sci Rep. 2013;3:3510.

22. Chen Y, Chen Q, McEachin RC, Cavalcoli JD, Yu X. H2a.B facilitates transcription elongation at methylated cpg loci. Genome Res. 2014;24:570–9.

23. Gu M, Naiyachit Y, Wood TJ, Millar CB. H2a.Z marks antisense promoters and has positive effects on antisense transcript levels in budding yeast. BMC Genomics. 2015;16:99.

24.   Jeronimo C, Watanabe S, Kaplan CD, Peterson CL, Robert F. The histone chaperones fact and spt6 restrict h2a.Z from intragenic locations. Mol Cell. 2015;58:1113–23.

25.   Weber CM, Ramachandran S, Henikoff S. Nucleosomes are context-specific, h2a.Z--modulated barriers to rna polymerase. Mol Cell. 2014;53:819–30.

26.   Wang J, Qiao M, He Q, Shi R, Loh SJ, Stanton LW, Wu M. Pluripotency activity of nanog requires biochemical stabilization by variant histone protein h2a.Z. Stem Cells. 2015;33:2126–34.

27.   Huh YH, Noh M, Burden FR, Chen JC, Winkler DA, Sherley JL. Sparse feature selection identifies h2a.Z as a novel, pattern-specific biomarker for asymmetrically self-renewing distributed stem cells. Stem Cell Res. 2015;14:144–54.

28.   Chen IY, Lypowy J, Pain J, Sayed D, Grinberg S, Alcendor RR, Sadoshima J, Abdellatif M. Histone h2a.Z is essential for cardiac myocyte hypertrophy but opposed by silent information regulator 2alpha. J Biol Chem. 2006;281:19369–77.

29.   Hartman PG, Chapman GE, Moss T, Bradbury EM. Studies on the role and mode of operation of the very-lysine-rich histone h1 in eukaryote chromatin. The three structural regions of the histone h1 molecule. Eur J Biochem/FEBS. 1977;77:45–51.

30.   Zhou BR, Jiang J, Feng H, Ghirlando R, Xiao TS, Bai Y. Structural mechanisms of nucleosome recognition by linker histones. Mol Cell. 2015;59:628–38.

31.   Vogler C, Huber C, Waldmann T, Ettig R, Braun L, Izzo A, Daujat S, Chassignet I, Lopez-Contreras AJ, Fernandez-Capetillo O, Dundr M, Rippe K, Langst G, Schneider R. Histone h2a c-terminus regulates chromatin dynamics, remodeling, and histone h1 binding. PLoS Genet. 2010;6:e1001234.

32.   Shukla MS, Syed SH, Goutte-Gattat D, Richard JL, Montel F, Hamiche A, Travers A, Faivre-Moskalenko C, Bednar J, Hayes JJ, Angelov D, Dimitrov S. The docking domain of histone h2a is required for h1 binding and rsc-mediated nucleosome remodeling. Nucleic Acids Res. 2011;39:2559–70.

33.   Szerlong HJ, Herman JA, Krause CM, DeLuca JG, Skoultchi A, Winger QA, Prenni JE, Hansen JC. Proteomic characterization of the nucleolar linker histone h1 interaction network. J Mol Biol. 2015;427:2056–71.

34.   Talasz H, Helliger W, Puschendorf B, Lindner H. In vivo phosphorylation of histone h1 variants during the cell cycle. Biochemistry. 1996;35:1761–7.

35. Lopez R, Sarg B, Lindner H, Bartolome S, Ponte I, Suau P, Roque A. Linker histone partial phosphorylation: effects on secondary structure and chromatin condensation. Nucleic Acids Res. 2015;43:4463–76.

36. Chen Y, Hoover ME, Dang X, Shomo AA, Guan X, Marshall AG, Freitas MA, Young NL. Quantitative mass spectrometry reveals that intact histone h1 phosphorylations are variant specific and exhibit single molecule hierarchical dependence. Mol Cell Proteomics MCP. 2015;15:813–33.

37. Fan Y, Nikitina T, Morin-Kensicki EM, Zhao J, Magnuson TR, Woodcock CL, Skoultchi AI. H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo. Mol Cell Biol. 2003;23:4559–72.

38. Zhang Y, Cooke M, Panjwani S, Cao K, Krauth B, Ho PY, Medrzycki M, Berhe DT, Pan C, McDevitt TC, Fan Y. Histone h1 depletion impairs embryonic stem cell differentiation. PLoS Genet. 2012;8:e1002691.

39. Vidali G, Boffa LC, Allfrey VG. Selective release of chromosomal proteins during limited dnaase 1 digestion of avian erythrocyte chromatin. Cell. 1977;12:409–15.

40. Fussner E, Strauss M, Djuric U, Li R, Ahmed K, Hart M, Ellis J, Bazett-Jones DP. Open and closed domains in the mouse genome are configured as 10-nm chromatin fibres. EMBO Rep. 2012;13:992–6.

41. Maeshima K, Imai R, Tamura S, Nozaki T. Chromatin as dynamic 10-nm fibers. Chromosoma. 2014;123:225–37.

42. Song F, Chen P, Sun D, Wang M, Dong L, Liang D, Xu RM, Zhu P, Li G. Cryo-em study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. Science. 2014;344:376–80.

43. Collepardo-Guevara R, Schlick T. Chromatin fiber polymorphism triggered by variations of DNA linker lengths. Proc Natl Acad Sci U S A. 2014;111:8061–6.

44. Cremer T, Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet. 2001;2:292–301.

45. Dekker J, Marti-Renom MA, Mirny LA. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet. 2013;14:390–403.

46. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485:376–80.

47. Ciabrelli F, Cavalli G. Chromatin-driven behavior of topologically associating domains. J Mol Biol. 2015;427:608–25.

48. Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, Vera DL, Wang Y, Hansen RS, Canfield TK, Thurman RE, Cheng Y, Gulsoy G, Dennis JH, Snyder MP, Stamatoyannopoulos JA, Taylor J, Hardison RC, Kahveci T, Ren B, Gilbert DM. Topologically associating domains are stable units of replication-timing regulation. Nature. 2014;515:402–5.

49. Lupianez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R, Santos-Simarro F, Gilbert-Dussardier B, Wittler L, Borschiwer M, Haas SA, Osterwalder M, Franke M, Timmermann B, Hecht J, Spielmann M, Visel A, Mundlos S. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell. 2015;161:1012–25.

50. Travis GH, Colavito-Shepanski M, Grunstein M. Extensive purification and characterization of chromatin-bound histone acetyltransferase from saccharomyces cerevisiae. J Biol Chem. 1984;259:14406–12.

51. Allfrey VG. In chromatin and chromosome structure. New York: Academic; 1977.

52. Zhao Y, Garcia BA. Comprehensive catalog of currently documented histone modifications. Cold Spring Harb Perspect Biol. 2015;7:a025064.

53. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. Nature. 2011;470:279–83.

54. Karlic R, Chung HR, Lasserre J, Vlahovicek K, Vingron M. Histone modification levels are predictive for gene expression. Proc Natl Acad Sci U S A. 2010;107:2926–31.

55. Gillette TG, Hill JA. Readers, writers, and erasers: chromatin as the whiteboard of heart disease. Circ Res. 2015;116:1245–53.

56. McKinsey TA, Zhang CL, Olson EN. Control of muscle development by dueling hats and hdacs. Curr Opin Genet Dev. 2001;11:497–504.

57. Awad S, Al-Haffar KM, Marashly Q, Quijada P, Kunhi M, Al-Yacoub N, Wade FS, Mohammed SF, Al-Dayel F, Sutherland G, Assiri A, Sussman M, Bers D, Al-Habeeb W, Poizat C. Control of histone h3 phosphorylation by camkiidelta in response to haemodynamic cardiac stress. J Pathol. 2015;235:606–18.

58. Sharma A, Nguyen H, Geng C, Hinman MN, Luo G, Lou H. Calcium-mediated histone modifications regulate alternative splicing in cardiomyocytes. Proc Natl Acad Sci U S A. 2014;111:E4920–8.

59. Blakeslee WW, Wysoczynski CL, Fritz KS, Nyborg JK, Churchill ME, McKinsey TA. Class i hdac inhibition stimulates cardiac protein sumoylation through a post-translational mechanism. Cell Signal. 2014;26:2912–20.

60. Papait R, Cattaneo P, Kunderfranco P, Greco C, Carullo P, Guffanti A, Vigano V, Stirparo GG, Latronico MV, Hasenfuss G, Chen J, Condorelli G. Genome-wide analysis of histone marks identifying an epigenetic signature of promoters and enhancers underlying cardiac hypertrophy. Proc Natl Acad Sci U S A. 2013;110:20164–9.

61. Capra JA. Extrapolating histone marks across developmental stages, tissues, and species: an enhancer prediction case study. BMC Genomics. 2015;16:104.

62. Heinig M, Colome-Tatche M, Taudt A, Rintisch C, Schafer S, Pravenec M, Hubner N, Vingron M, Johannes F. Histonehmm: differential analysis of histone modifications with broad genomic footprints. BMC Bioinf. 2015;16:60.

63. Rintisch C, Heinig M, Bauerfeind A, Schafer S, Mieth C, Patone G, Hummel O, Chen W, Cook S, Cuppen E, Colome-Tatche M, Johannes F, Jansen RC, Neil H, Werner M, Pravenec M, Vingron M, Hubner N. Natural variation of histone modification and its impact on gene expression in the rat genome. Genome Res. 2014;24:942–53.

64. Whitehouse I, Flaus A, Cairns BR, White MF, Workman JL, Owen-Hughes T. Nucleosome mobilization catalysed by the yeast swi/snf complex. Nature. 1999;400:784–7.

65. Dechassa ML, Sabri A, Pondugula S, Kassabov SR, Chatterjee N, Kladde MP, Bartholomew B. Swi/snf has intrinsic nucleosome disassembly activity that is dependent on adjacent nucleosomes. Mol Cell. 2010;38:590–602.

66. Mizuguchi G, Shen X, Landry J, Wu WH, Sen S, Wu C. Atp-driven exchange of histone h2az variant catalyzed by swr1 chromatin remodeling complex. Science. 2004;303:343–8.

67. Payne S, Burney MJ, McCue K, Popal N, Davidson SM, Anderson RH, Scambler PJ. A critical role for the chromatin remodeller chd7 in anterior mesoderm during cardiovascular development. Dev Biol. 2015;405:82–95.

68. Hang CT, Yang J, Han P, Cheng HL, Shang C, Ashley E, Zhou B, Chang CP. Chromatin regulation by brg1 underlies heart muscle development and disease. Nature. 2010;466:62–7.

69. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases dnmt3a and dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999;99:247–57.

70. Deaton AM, Bird A. Cpg islands and the regulation of transcription. Genes Dev. 2011;25:1010–22.

71. Jones PA. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012;13:484–92.

72. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by mll partner tet1. Science. 2009;324:930–5.

73. Lev Maor G, Yearim A, Ast G. The alternative role of DNA methylation in splicing regulation. Trends Genet TIG. 2015;31:274–80.

74. Serra-Juhe C, Cusco I, Homs A, Flores R, Toran N, Perez-Jurado LA. DNA methylation abnormalities in congenital heart disease. Epigenetics. 2015;10:167–77.

75. Sheng W, Qian Y, Zhang P, Wu Y, Wang H, Ma X, Chen L, Ma D, Huang G. Association of promoter methylation statuses of congenital heart defect candidate genes with tetralogy of fallot. J Transl Med. 2014;12:31.

76. Haas J, Frese KS, Park YJ, Keller A, Vogel B, Lindroth AM, Weichenhan D, Franke J, Fischer S, Bauer A, Marquart S, Sedaghat-Hamedani F, Kayvanpour E, Kohler D, Wolf NM, Hassel S, Nietsch R, Wieland T, Ehlermann P, Schultz JH, Dosch A, Mereles D, Hardt S, Backs J, Hoheisel JD, Plass C, Katus HA, Meder B. Alterations in cardiac DNA methylation in human dilated cardiomyopathy. EMBO Mol Med. 2013;5:413–29.

77. Movassagh M, Choy MK, Goddard M, Bennett MR, Down TA, Foo RS. Differential DNA methylation correlates with differential expression of angiogenic factors in human heart failure. PLoS One. 2010;5:e8564.

78. Chamberlain AA, Lin M, Lister RL, Maslov AA, Wang Y, Suzuki M, Wu B, Greally JM, Zheng D, Zhou B. DNA methylation is developmentally regulated for genes essential for cardiogenesis. J Am Heart Assoc. 2014;3:e000976.

79. Hon GC, Rajagopal N, Shen Y, McCleary DF, Yue F, Dang MD, Ren B. Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. Nat Genet. 2013;45:1198–206.

80. Gilsbach R, Preissl S, Gruning BA, Schnick T, Burger L, Benes V, Wurch A, Bonisch U, Gunther S, Backofen R, Fleischmann BK, Schubeler D, Hein L. Dynamic DNA methylation orchestrates cardiomyocyte development, maturation and disease. Nat Commun. 2014;5:5288.

81. Paradis A, Xiao D, Zhou J, Zhang L. Endothelin-1 promotes cardiomyocyte terminal differentiation in the developing heart via heightened DNA methylation. Int J Med Sci. 2014;11:373–80.

82. Xiao D, Dasgupta C, Chen M, Zhang K, Buchholz J, Xu Z, Zhang L. Inhibition of DNA methylation reverses norepinephrine-induced cardiac hypertrophy in rats. Cardiovasc Res. 2014;101:373–82.

83. Cech TR, Steitz JA. The noncoding rna revolution-trashing old rules to forge new ones. Cell. 2014;157:77–94.

84. Su H, Trombly MI, Chen J, Wang X. Essential and overlapping functions for mammalian argonautes in microrna silencing. Genes Dev. 2009;23:304–17.

85. Li LC. Chromatin remodeling by the small rna machinery in mammalian cells. Epigenetics. 2014;9:45–52.

86. Zhang B, Arun G, Mao YS, Lazar Z, Hung G, Bhattacharjee G, Xiao X, Booth CJ, Wu J, Zhang C, Spector DL. The lncrna malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. Cell Rep. 2012;2:111–23.

87. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, Wang Y, Brzoska P, Kong B, Li R, West RB, van de Vijver MJ, Sukumar S, Chang HY. Long non-coding rna hotair reprograms chromatin state to promote cancer metastasis. Nature. 2010;464:1071–6.

88. Moazed D. Small rnas in transcriptional gene silencing and genome defence. Nature. 2009;457:413–20.

89. Lam MT, Li W, Rosenfeld MG, Glass CK. Enhancer rnas and regulated transcriptional programs. Trends Biochem Sci. 2014;39:170–82.

90. Li L, Chang HY. Physiological roles of long noncoding rnas: insight from knockout mice. Trends Cell Biol. 2014;24:594–602.

91. Klattenhoff CA, Scheuermann JC, Surface LE, Bradley RK, Fields PA, Steinhauser ML, Ding H, Butty VL, Torrey L, Haas S, Abo R, Tabebordbar M, Lee RT, Burge CB, Boyer LA. Braveheart, a long noncoding rna required for cardiovascular lineage commitment. Cell. 2013;152:570–83.

92. Grote P, Wittler L, Hendrix D, Koch F, Wahrisch S, Beisaw A, Macura K, Blass G, Kellis M, Werber M, Herrmann BG. The tissue-specific lncrna fendrr is an essential regulator of heart and body wall development in the mouse. Dev Cell. 2013;24:206–14.

93. Han P, Li W, Lin CH, Yang J, Shang C, Nurnberg ST, Jin KK, Xu W, Lin CY, Lin CJ, Xiong Y, Chien HC, Zhou B, Ashley E, Bernstein D, Chen PS, Chen HS, Quertermous T, Chang CP. A long noncoding rna protects the heart from pathological hypertrophy. Nature. 2014;514:102–6.

94. Ishii N, Ozaki K, Sato H, Mizuno H, Saito S, Takahashi A, Miyamoto Y, Ikegawa S, Kamatani N, Hori M, Saito S, Nakamura Y, Tanaka T. Identification of a novel non-coding rna, miat, that confers risk of myocardial infarction. J Hum Genet. 2006;51:1087–99.

95. Boon RA, Dimmeler S. Micrornas in myocardial infarction. Nat Rev Cardiol. 2015;12:135–42.

96. Dorn 2nd GW. Decoding the cardiac message: the 2011 thomas w. Smith memorial lecture. Circ Res. 2012;110:755–63.

97. Jeck WR, Sharpless NE. Detecting and characterizing circular rnas. Nat Biotechnol. 2014;32:453–61.

98. Burd CE, Jeck WR, Liu Y, Sanoff HK, Wang Z, Sharpless NE. Expression of linear and novel circular forms of an ink4/arf-associated non-coding rna correlates with atherosclerosis risk. PLoS Genet. 2010;6:e1001233.

99. Szabo L, Morey R, Palpant NJ, Wang PL, Afari N, Jiang C, Parast MM, Murry CE, Laurent LC, Salzman J. Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular rna during human fetal development. Genome Biol. 2015;16:126.

100. Boeckel JN, Jae N, Heumuller AW, Chen W, Boon RA, Stellos K, Zeiher AM, John D, Uchida S, Dimmeler S. Identification and characterization of hypoxia-regulated endothelial circular rna. Circ Res. 2015;117:884–90.

101. Chen PB, Zhu LJ, Hainer SJ, McCannell KN, Fazzio TG. Unbiased chromatin accessibility profiling by red-seq uncovers unique features of nucleosome variants in vivo. BMC Genomics. 2014;15:1104.

102. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, Gribnau J, Barillot E, Bluthgen N, Dekker J, Heard E. Spatial partitioning of the regulatory landscape of the x-inactivation centre. Nature. 2012;485:381–5.

103. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EG, Huang PY, Welboren WJ, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY, Wansa KD, Zhao B, Lim KS, Leow SC, Yow JS, Joseph R, Li H, Desai KV, Thomsen JS, Lee YK, Karuturi RK, Herve T, Bourque G, Stunnenberg HG, Ruan X, Cacheux-Rataboul V, Sung WK, Liu ET,

Wei CL, Cheung E, Ruan Y. An oestrogen-receptor-alpha-bound human chromatin interactome. Nature. 2009;462:58–64.

104. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013;10:1213–8.

105. Ishii H, Kadonaga JT, Ren B. Mpe-seq, a new method for the genome-wide analysis of chromatin structure. Proc Natl Acad Sci U S A. 2015;112:E3457–65.

106. Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, Bernstein BE, Bickel P, Brown JB, Cayting P, Chen Y, DeSalvo G, Epstein C, Fisher-Aylor KI, Euskirchen G, Gerstein M, Gertz J, Hartemink AJ, Hoffman MM, Iyer VR, Jung YL, Karmakar S, Kellis M, Kharchenko PV, Li Q, Liu T, Liu XS, Ma L, Milosavljevic A, Myers RM, Park PJ, Pazin MJ, Perry MD, Raha D, Reddy TE, Rozowsky J, Shoresh N, Sidow A, Slattery M, Stamatoyannopoulos JA, Tolstorukov MY, White KP, Xi S, Farnham PJ, Lieb JD, Wold BJ, Snyder M. Chip-seq guidelines and practices of the encode and modencode consortia. Genome Res. 2012;22:1813–31.

107. Urich MA, Nery JR, Lister R, Schmitz RJ, Ecker JR. Methylc-seq library preparation for base-resolution whole-genome bisulfite sequencing. Nat Protoc. 2015;10:475–83.

108. Booth MJ, Branco MR, Ficz G, Oxley D, Krueger F, Reik W, Balasubramanian S. Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science. 2012;336:934–7.

109. Booth MJ, Marsico G, Bachman M, Beraldi D, Balasubramanian S. Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. Nat Chem. 2014;6:435–40.

110. Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. Genomic maps of long noncoding rna occupancy reveal principles of rna-chromatin interactions. Mol Cell. 2011;44:667–78.

111. Churchman LS, Weissman JS. Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature. 2011;469:368–73.

112. Churko JM, Mantalas GL, Snyder MP, Wu JC. Overview of high throughput sequencing technologies to elucidate molecular pathways in cardiovascular diseases. Circ Res. 2013;112:1613–23.

113. Li Y, Klena NT, Gabriel GC, Liu X, Kim AJ, Lemke K, Chen Y, Chatterjee B, Devine W, Damerla RR, Chang C, Yagi H, San Agustin JT, Thahir M, Anderton S, Lawhead C, Vescovi A, Pratt H, Morgan J, Haynes L, Smith CL, Eppig JT, Reinholdt L, Francis R, Leatherbury L, Ganapathiraju MK, Tobita K, Pazour GJ,

Lo CW. Global genetic analysis in mice unveils central role for cilia in congenital heart disease. Nature. 2015;521:520–4.

114. Do R, Stitziel NO, Won HH, Jorgensen AB, Duga S, Angelica Merlini P, Kiezun A, Farrall M, Goel A, Zuk O, Guella I, Asselta R, Lange LA, Peloso GM, Auer PL, Project NES, Girelli D, Martinelli N, Farlow DN, DePristo MA, Roberts R, Stewart AF, Saleheen D, Danesh J, Epstein SE, Sivapalaratnam S, Hovingh GK, Kastelein JJ, Samani NJ, Schunkert H, Erdmann J, Shah SH, Kraus WE, Davies R, Nikpay M, Johansen CT, Wang J, Hegele RA, Hechter E, Marz W, Kleber ME, Huang J, Johnson AD, Li M, Burke GL, Gross M, Liu Y, Assimes TL, Heiss G, Lange EM, Folsom AR, Taylor HA, Olivieri O, Hamsten A, Clarke R, Reilly DF, Yin W, Rivas MA, Donnelly P, Rossouw JE, Psaty BM, Herrington DM, Wilson JG, Rich SS, Bamshad MJ, Tracy RP, Cupples LA, Rader DJ, Reilly MP, Spertus JA, Cresci S, Hartiala J, Tang WH, Hazen SL, Allayee H, Reiner AP, Carlson CS, Kooperberg C, Jackson RD, Boerwinkle E, Lander ES, Schwartz SM, Siscovick DS, McPherson R, Tybjaerg-Hansen A, Abecasis GR, Watkins H, Nickerson DA, Ardissino D, Sunyaev SR, O'Donnell CJ, Altshuler D, Gabriel S, Kathiresan S. Exome sequencing identifies rare ldlr and apoa5 alleles conferring risk for myocardial infarction. Nature. 2015;518:102–6.

115. Myocardial Infarction Genetics Consortium I, Stitziel NO, Won HH, Morrison AC, Peloso GM, Do R, Lange LA, Fontanillas P, Gupta N, Duga S, Goel A, Farrall M, Saleheen D, Ferrario P, Konig I, Asselta R, Merlini PA, Marziliano N, Notarangelo MF, Schick U, Auer P, Assimes TL, Reilly M, Wilensky R, Rader DJ, Hovingh GK, Meitinger T, Kessler T, Kastrati A, Laugwitz KL, Siscovick D, Rotter JI, Hazen SL, Tracy R, Cresci S, Spertus J, Jackson R, Schwartz SM, Natarajan P, Crosby J, Muzny D, Ballantyne C, Rich SS, O'Donnell CJ, Abecasis G, Sunyaev S, Nickerson DA, Buring JE, Ridker PM, Chasman DI, Austin E, Ye Z, Kullo IJ, Weeke PE, Shaffer CM, Bastarache LA, Denny JC, Roden DM, Palmer C, Deloukas P, Lin DY, Tang ZZ, Erdmann J, Schunkert H, Danesh J, Marrugat J, Elosua R, Ardissino D, McPherson R, Watkins H, Reiner AP, Wilson JG, Altshuler D, Gibbs RA, Lander ES, Boerwinkle E, Gabriel S, Kathiresan S. Inactivating mutations in npc1l1 and protection from coronary heart disease. N Engl J Med. 2014;371:2072–82.

116. Tao H, Yang JJ, Shi KH, Deng ZY, Li J. DNA methylation in cardiac fibrosis: new advances and perspectives. Toxicology. 2014;323:125–9.

117. Zhao JY, Yang XY, Shi KH, Sun SN, Hou J, Ye ZZ, Wang J, Duan WY, Qiao B, Chen YJ, Shen HB, Huang GY, Jin L, Wang HY. A functional variant in the cystathionine beta-synthase gene promoter significantly reduces congenital heart disease susceptibility in a han chinese population. Cell Res. 2012;23:242–53.

118. Sanchez-Castro M, Gordon CT, Petit F, Nord AS, Callier P, Andrieux J, Guerin P, Pichon O, David A, Abadie V, Bonnet D, Visel A, Pennacchio LA, Amiel J,

Lyonnet S, Le Caignec C. Congenital heart defects in patients with deletions upstream of sox9. Hum Mutat. 2013;34:1628–31.

119. Morikawa Y, Zhang M, Heallen T, Leach J, Tao G, Xiao Y, Bai Y, Li W, Willerson JT, Martin JF. Actin cytoskeletal remodeling with protrusion formation is essential for heart regeneration in hippo-deficient mice. Sci Signal. 2015;8:ra41.

120. Heallen T, Zhang M, Wang J, Bonilla-Claudio M, Klysik E, Johnson RL, Martin JF. Hippo pathway inhibits wnt signaling to restrain cardiomyocyte proliferation and heart size. Science. 2011;332:458–61.

121. Huang GN, Thatcher JE, McAnally J, Kong Y, Qi X, Tan W, DiMaio JM, Amatruda JF, Gerard RD, Hill JA, Bassel-Duby R, Olson EN. C/ebp transcription factors mediate epicardial activation during heart development and injury. Science. 2012;338:1599–603.

122. Takeuchi JK, Lou X, Alexander JM, Sugizaki H, Delgado-Olguin P, Holloway AK, Mori AD, Wylie JN, Munson C, Zhu Y, Zhou YQ, Yeh RF, Henkelman RM, Harvey RP, Metzger D, Chambon P, Stainier DY, Pollard KS, Scott IC, Bruneau BG. Chromatin remodelling complex dosage modulates transcription factor function in heart development. Nat Commun. 2011;2:187.

123. Lei I, Gao X, Sham MH, Wang Z. Swi/snf protein component baf250a regulates cardiac progenitor cell differentiation by modulating chromatin accessibility during second heart field development. J Biol Chem. 2012;287:24255–62.

124. He A, Ma Q, Cao J, von Gise A, Zhou P, Xie H, Zhang B, Hsing M, Christodoulou DC, Cahan P, Daley GQ, Kong SW, Orkin SH, Seidman CE, Seidman JG, Pu WT. Polycomb repressive complex 2 regulates normal development of the mouse heart. Circ Res. 2012;110:406–15.

125. Anand P, Brown JD, Lin CY, Qi J, Zhang R, Artero PC, Alaiti MA, Bullard J, Alazem K, Margulies KB, Cappola TP, Lemieux M, Plutzky J, Bradner JE, Haldar SM. Bet bromodomains mediate transcriptional pause release in heart failure. Cell. 2013;154:569–82.

126. He A, Kong SW, Ma Q, Pu WT. Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. Proc Natl Acad Sci U S A. 2011;108:5632–7.

127. Mitchell-Jordan S, Chen H, Franklin S, Stefani E, Bentolila LA, Vondriska TM. Features of endogenous cardiomyocyte chromatin revealed by super-resolution sted microscopy. J Mol Cell Cardiol. 2012;53:552–8.

128. Wood AJ, Lo TW, Zeitler B, Pickle CS, Ralston EJ, Lee AH, Amora R, Miller JC, Leung E, Meng X, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Meyer BJ.

Targeted genome editing across species using zfns and talens. Science. 2011;333:307.

129. Hsu PD, Lander ES, Zhang F. Development and applications of crispr-cas9 for genome engineering. Cell. 2014;157:1262–78.

130. Ghamari A, van de Corput MP, Thongjuea S, van Cappellen WA, van Ijcken W, van Haren J, Soler E, Eick D, Lenhard B, Grosveld FG. In vivo live imaging of rna polymerase ii transcription factories in primary cells. Genes Dev. 2013;27:767–77.

131. Wang Y, Maharana S, Wang MD, Shivashankar GV. Super-resolution microscopy reveals decondensed chromatin structure at transcription sites. Sci Rep. 2014;4:4477.

132. Ricci MA, Manzo C, Garcia-Parajo MF, Lakadamyali M, Cosma MP. Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. Cell. 2015;160:1145–58.

133. Kohl T, Westphal V, Hell SW, Lehnart SE. Superresolution microscopy in heart – cardiac nanoscopy. J Mol Cell Cardiol. 2013;58:13–21.

134. Bhat S, Ohn J, Liebling M. Motion-based structure separation for label-free, high-speed, 3d cardiac microscopy. IEEE Trans Image Process Publ IEEE Signal Process Soc. 2012;21:3638–47.

135. Lee S, Vinegoni C, Feruglio PF, Fexon L, Gorbatov R, Pivoravov M, Sbarbati A, Nahrendorf M, Weissleder R. Real-time in vivo imaging of the beating mouse heart at microscopic resolution. Nat Commun. 2012;3:1054.

136. Jung K, Kim P, Leuschner F, Gorbatov R, Kim JK, Ueno T, Nahrendorf M, Yun SH. Endoscopic time-lapse imaging of immune cells in infarcted mouse hearts. Circ Res. 2013;112:891–9.

137. Nance ME, Whitfield JT, Zhu Y, Gibson AK, Hanft LM, Campbell KS, Meininger GA, McDonald KS, Segal SS, Domeier TL. Attenuated sarcomere lengthening of the aged murine left ventricle observed using two-photon fluorescence microscopy. Am J Physiol Heart Circ Physiol. 2015;309:H918–25.

138. Sewell-Loftin MK, Brown CB, Baldwin HS, Merryman WD. A novel technique for quantifying mouse heart valve leaflet stiffness with atomic force microscopy. J Heart Valve Dis. 2012;21:513–20.

139. Pauza DH, Rysevaite K, Inokaitis H, Jokubauskas M, Pauza AG, Brack KE, Pauziene N. Innervation of sinoatrial nodal cardiomyocytes in mouse. A combined approach using immunofluorescent and electron microscopy. J Mol Cell Cardiol. 2014;75:188–97.

140. Leo-Macias A, Liang FX, Delmar M. Ultrastructure of the intercellular space in adult murine ventricle revealed by quantitative tomographic electron microscopy. Cardiovasc Res. 2015;107:442–52.

141. Fujioka H, Tandler B, Hoppel CL. Mitochondrial division in rat cardiomyocytes: an electron microscope study. Anat Rec. 2012;295:1455–61.

142. Gherghiceanu M, Popescu LM. Heterocellular communication in the heart: electron tomography of telocyte-myocyte junctions. J Cell Mol Med. 2011;15:1005–11.

Chapter 1: The Chromatin-binding Protein Smyd1 Restricts Adult Mammalian Heart Growth

Sarah Franklin, Todd Kimball, Tara L. Rasmussen, Manuel Rosa-Garrido, Haodong Chen, Tam Tran, Mickey R. Miller, Ricardo Gray, Shanxi Jiang, Shuxun Ren, Yibin Wang, Haley O. Tucker, and Thomas M. Vondriska

Abstract

All terminally differentiated organs face two challenges, maintaining their cellular identity and restricting organ size. The molecular mechanisms responsible for these decisions are of critical importance to organismal development, and perturbations in their normal balance can lead to disease. A hallmark of heart failure, a condition affecting millions of people worldwide, is hypertrophic growth of cardiomyocytes. The various forms of heart failure in human and animal models share conserved transcriptome remodeling events that lead to expression of genes normally silenced in the healthy adult heart. However, the chromatin remodeling events that maintain cell and organ size are incompletely understood; insights into these mechanisms could provide new targets for heart failure therapy. Using a quantitative proteomics approach to identify muscle-specific chromatin regulators in a mouse model of hypertrophy and

heart failure, we identified upregulation of the histone methyltransferase Smyd1 during disease. Inducible loss-of-function studies in vivo demonstrate that Smyd1 is responsible for restricting growth in the adult heart, with its absence leading to cellular hypertrophy, organ remodeling, and fulminate heart failure. Molecular studies reveal Smyd1 to be a muscle-specific regulator of gene expression and indicate that Smyd1 modulates expression of gene isoforms whose expression is associated with cardiac pathology. Importantly, activation of Smyd1 can prevent pathological cell growth. These findings have basic implications for our understanding of cardiac pathologies and open new avenues to the treatment of cardiac hypertrophy and failure by modulating Smyd1.

Smyd1; epigenetics; heart failure; cardiac hypertrophy; histone methyltransferase

NEW & NOTEWORTHY

This study is the first to demonstrate that loss of a muscle-specific chromatin-binding protein, Smyd1, is sufficient to induce cardiac hypertrophy and failure. Moreover, the findings demonstrate that augmentation of Smyd1 levels can block hypertrophy in cell models. These studies may support novel strategies for cardiac-targeted epigenetic therapy.

Heart failure is a particularly nefarious result of many different forms of cardiovascular disease and has a massive human health burden; >5 million Americans have heart failure (a number projected to increase to >8 million by 2030) with ~825,000 new cases annually. Diabetes mellitus, hypertension, heart attacks, and atherosclerosis

all predispose patients to heart failure, and, although pharmacological treatments are available, outcomes are dismal with a 50% mortality rate 5 yr after diagnosis (13). Clearly, novel approaches for treatment of heart failure are desperately needed.

During development, the mammalian heart expands the cardiac myocyte pool through hyperplasia. Soon after birth, however, myocytes (the contractile unit of the heart) undergo mitosis without cytokinesis and then exit the cell cycle (39). Like most terminally differentiated cells, adult cardiac myocytes lack the ability to proliferate and, in response to environmental stress, adapt via hypertrophic growth (9, 47). This hypertrophic response can be initially beneficial to maintain the workload of the heart following injuries such as myocardial infarction, but animal models and human studies have consistently shown that hypertrophic growth in the heart is a precursor to cellular and organ failure (18). It has been well established that heart failure involves activation of a transcriptional network that is conserved across animal models and humans. Specifically, genes associated with normal muscle development, but silenced in the healthy adult heart, become reactivated (40). Development and disease are regulated by a host of myocyte-specific transcription factors that regulate cardiac genes (37, 44). Manipulating this network, therefore, has significant therapeutic potential; however, the molecular entities regulating these genomic targets at the level of chromatin have been elusive.

Recent studies have provided new insights into the epigenetic cues that determine distinct developmental transitions from pluripotency to mature cardiac myocyte (38, 49), but the landscape in the adult heart is more complex. Previous studies have demonstrated that histone deacetylases play a powerful role in controlling

40

cardiac growth (36, 52). Subsequent studies on chromatin-modifying enzymes have implicated specific classes and modifications of histone deacetylases (3, 23), acetyltransferases (50), chromatin-remodeling complexes (15), and histone mark readers (4) in heart failure. A fundamental gap in our understanding of the mammalian heart relates to how cardiac phenotype is established and maintained according to cardiac myocyte-specific chromatin modifiers. The absence of such insights also limits therapeutic targeting of cardiac chromatin, insofar as the heart is comprised of many distinct cell types, most if not all of which express the epigenetic modifiers explored in these previous studies.

To identify chromatin modifiers responsible for maintenance of the adult cardiac myocyte transcriptome (and phenotype), we used quantitative proteomics to measure proteins associated with chromatin in the mouse heart and to detect changes occurring following heart failure induced by pressure overload. We then filtered the proteins according to cell type specificity and alteration with disease, examining only those that increased with heart failure. Among proteins in this group, only one was muscle specific, the SET and MYND domain containing histone methyltransferase, Smyd1. Previous studies on Smyd1 (also known as Bop) have demonstrated its role in early cardiac development (14), but its role in the adult heart was unknown. We generated mice with Smyd1 specifically deleted in an inducible manner in the adult cardiac myocyte. These animals develop heart failure following activation of a gene expression profile normally repressed by Smyd1 and associated with cardiac remodeling and dysfunction. We demonstrate that Smyd1 acts in part via transcriptional repression and can prevent

41

cardiac hypertrophy at the cellular level. These findings implicate Smyd1 histone methyltransferase as a novel muscle-specific molecular target for heart failure.

METHODS

Cardiac-specific inducible Smyd1 knockout mouse.

All protocols involving animals conform to the NIH *Guide for the Care and Use of Laboratory Animals* and were approved by the UCLA Animal Research Committee. The endogenous Smyd1 sequence was obtained from a Lambda FixII Vector 129SV Mouse Genomic Library (Stratagene). LoxP sites were introduced, flanking the second and third exons of *smyd1* followed by a neomycin cassette flanked by a third LoxP site, using a targeting construct designed in the Osdupdel vector (gift from William A. Kuziel). This targeting vector was electroporated into 129S6 embryonic stem (ES) cells, and surviving clones, selected with gancyclovir and G418, were screened for homologous recombination by Southern analysis by digestion with either Bgl II and Sal I or Bgl II and Kpn I and hybridization with probe 2 and probe 3, respectively. ES clones showing correct targeting in both arms were injected into C57BL/6 blastocysts to create chimeric mice, which were then mated to C57BL/6 females to create a germline knockin (*Smyd1KI*). *Smyd1KI* mice, B6.129 hybrids, were crossed to ubiquitously expressing *EIIa-Cre* mice, B6.FVB hybrids, and progeny were screened for deletion of the neomycin cassette by Southern analysis and backcrossed to remove *EIIa-Cre*. Resulting *Smyd1^{flox}* mice, lacking *EIIa-Cre*, were crossed with α-MHC-MerCreMer mice purchased from Jackson Laboratory (cat. no. 005657). To induce recombination, mice were fed a diet containing tamoxifen, 0.4 mg/g of chow diet (Harlan, cat. no. TD.07262).

All mice used in this study were males 8–10 wk of age upon beginning tamoxifen chow feeding. To examine DNA synthesis and cell cycle progression, mice were injected with three doses of 5-bromo-1-(2-deoxy-β-D-ribofuranosyl) uracil (BrdU; 50 mg/kg body wt) at 12-h intervals and killed 2–4 h after the last injection.

Echocardiography.

Echocardiography (ECHO) was performed on a Vevo 2100 (Visual Sonics) to determine cardiac parameters in live mice as described (11, 32), including the following indices: left ventricular size (end-diastolic and end-systolic dimension), wall thickness (intraventricular septum and posterior wall thickness), ventricular mass, and ejection fraction.

Histology.

At the time of death, mice were anesthetized, and the heart was arrested in diastole with an intracardiac injection of 2,3-butanedione monoxime (10 mM). Whole hearts were rapidly excised from the animals fixed in 4% paraformaldehyde and then embedded in paraffin. Hearts were then sectioned at 4 μm and placed on slides. Small intestine was also harvested as a control for BrdU incorporation studies, as this tissue has a high cellular turnover rate.

Heart sections were incubated with Texas red-X-conjugated wheat germ agglutinin (Invitrogen, 1:100) for 90 min and assessed for cell size. Hematoxylin and eosin (H and E) staining and Masson's trichrome (Sigma) staining were performed according to the manufacturer's protocols. Tissue sections were incubated with primary

antibodies to BrdU (ab6326, Abcam; secondary: A11006, Invitrogen) and α-actinin (A7811, Sigma; secondary: A11004, Invitrogen) and mounted with Vectashield containing DAPI (H-1200, Vector Laboratories). Tissue sections were visualized and imaged on either a Nikon eclipse TE2000-U microscope with SPOT software (Diagnostic Instruments) or a Nikon A1R MP multiphoton confocal microscope.

Cell fractionation.

Murine heart tissue and isolated neonatal rat ventricular myocytes (NRVMs) were subject to cellular fractionation as previously described to examine compartmental-specific changes in protein abundance and localization (12, 33).

Microarrays.

For microarray analysis, heart tissue was removed from the left and right ventricles of Smyd1 knockout (KO) animals (or normal diet-fed control mice) 2 or 9 wk after tamoxifen treatment. Total RNA was isolated using TRIzol (Invitrogen) according to the manufacturer's protocol. RNA was analyzed for genome-wide expression analysis using an Illumina mouse bead chip (Mouse Ref 8 v. 2.0). All total RNA samples were quantified using a Ribogreen fluorescent assay and normalized to 10 ng/µl before amplification. Amplified and labeled cRNA was produced from 100 ng of each sample using the Illumina-specific version of the Ambion TotalPrep 96 kit (cat. no. 4393543). After a second Ribogreen quantification and normalization step, amplified and labeled cRNA was hybridized overnight at 58°C to the expression arrays. Washing and signal development were performed with the aid of a SciGene model 650c microarray

processor (LittleDipper). Chips were scanned on an Illumina iScan confocal scanner under standard parameters. Expression data was extracted and compiled using BeadStudio software (Illumina). Samples were analyzed in triplicate for each group, and data for each gene were averaged between the three samples. Data were subjected to background subtraction and quantile normalization.

Isolated rat cardiomyocytes.

NRVMs were obtained by enzymatic dissociation from 1-day-old litters and plated in DMEM media (Invitrogen, no. 11965) containing 1% penicillin, 1% streptomycin, 1% insulin-transferrin-sodium selenite supplement, and 10% fetal bovine serum for the first 24 h, after which the cells are cultured in serum-free media. NRVMs were treated with adenovirus expressing either FLAG-tagged mouse Smyd1(a), Smyd1(b), or an empty virus (control). To induce hypertrophy in isolated NRVMs, cells were treated with either isoproterenol (ISO, 1 µM) or phenylephrine (PE, 10 µM) for 48 h. For cell size analysis, NRVMs were fixed with paraformaldehyde and stained with Alexa Fluor 488 phalloidin (A12379, Invitrogen) according to the manufacturer's protocol and imaged on a Nikon eclipse TE2000-U microscope.

Electrophoresis and Western blotting.

Proteins were separated by standard SDS-PAGE using Laemmli buffer. For Western blotting, proteins were transferred to nitrocellulose, membranes blocked with milk, and protein signals detected by enzyme-linked chemiluminescence (GE Biosciences). Ponceau staining of membranes was used to confirm transfer and protein

loading. Antibodies used in this study are as follows, including source and dilution: histone H3 (Abcam, ab1791; 1:10,000 dilution); histone H3-trimethylated-K9 (Abcam, ab8898; 1:500 dilution); histone H3-trimethylated K4 (Abcam, ab8580; 1:300 dilution); histone H4 (Abcam, ab10158; 1:1,000); histone H4-trimethylated K20 (Abcam, ab9053; 1:300 dilution); β-actin (Sigma, A1978; 1:1,000 dilution); Gapdh (Santa Cruz Biotechnology, sc-20357; 1:1,000 dilution); FLAG (Sigma, F1804; 1:1,000 dilution); Smyd1 (Abcam, 32482; 1:1,000 dilution); Smyd1 (Santa Cruz Biotechnology, sc-79080; 1:1,000 dilution); p53 (Santa Cruz Biotechnology, sc-1313; 1:500 dilution); Smyd2 (Abcam, ab38821; 1:1,000 dilution); HSP90 (Abcam, ab13494; 1:300 dilution); HSP90 (Cell Signaling, no. 4874; 1:1,000 dilution).

Quantitative real-time PCR analysis.

Total RNA was isolated from the left ventricle of the heart or from cultured NRVMs using TRIzol (Invitrogen) according to the manufacturer's protocol. Total RNA was transcribed using SuperScript First-Strand Synthesis system for RT-PCR (Invitrogen) according to the manufacturer's protocol to produce cDNA. cDNA transcripts were amplified on the iCycler iQ real-time PCR detection system with iQ SYBR Green Supermix (Bio-Rad). Expression levels were analyzed using the iQ5 Optical Systems software v2.0 and normalized against GAPDH by subtracting the mean cycle number for each experimental group from the mean cycle number for GAPDH from the same group. Fold change was calculated using the ΔΔCt method.

Luciferase reporter assay.

HeLa cells were grown in DMEM (Corning Cellgro) supplemented with 10% FBS and 1% penicillin/streptomycin on white 96-well plates. At 70% confluence, the cells were infected with either Ad-CMV-Null (empty virus) or Ad-Smyd1A-Flag adenovirus at a multiplicity of infection (MOI) of 100 in serum-free media. After 24 h, the cells were transfected with 75 ng of human TGF-β3, Nppa, negative control (scrambled), or positive control (actin) luciferase reporter construct (SwitchGear Genomics) using FuGene HD (SwitchGear Genomics) at a 6:1 ratio to DNA and incubated for 48 h. Luciferase activity was assayed using the LightSwitch Luciferase assay reagent (SwitchGear Genomics) according to the manufacturer's instructions and measured using a BioTek Synergy Neo HTS Multi-Mode microplate reader. This assay was also performed in NRVM cells; however, despite the detection of signal from the actin positive control construct, no signal was observed for TGF-β3 or Nppa. The promoter regions for these genes in human and rat contain significant sequence variability and may be the reason for the absence of signal in rat cells.

ChIP-PCR.

NRVMs (6×10^6) infected with adenovirus-expressing Smyd1a-FLAG or an empty vector control (in the presence or absence of PE treatment, 48 h) were fixed in 1% formaldehyde, lysed in membrane extraction buffer (Magnetic ChIP Kit), and sonicated using a Sonic Dismembrator (Fisher Scientific), leading to fragments between 300 and 1,000 bp (for endogenous ChIP experiments, wild-type or Smyd1 KO heart lysates were used). ChIP was performed using a commercially available Pierce Magnetic ChIP Kit (Thermo, 26157) according to manufacturer's instructions. DNA-bound protein was

immunoprecipitated using an anti-FLAG (Sigma, F1804), anti-histone H3 trimethylated K4 (Abcam, ab8580), and anti-IgG (Santa Cruz Biotechnology, 2025) as a negative control. The DNA recovered was analyzed by quantitative real-time PCR using primer sets that amplified the promoter region of Tgf-β3 or atrial natriuretic factor (ANF) (Nppa). PCR was performed in duplicate with equal amounts of specific antibody-immunoprecipitated sample, control (IgG) and Input. Values were normalized to input measurements, and enrichment was calculated using the ΔΔCt method.

Bioinformatic analysis.

The heatmap in Fig. 5*A* was generated using the function "heatmap.2" from gplots package (2.11.3) in R (3.0.1). The distance/dissimilarity between genes was calculated using Euclidean distance, as default. To achieve a better separation between upregulated expression (red) and downregulated expression (green), the single linkage method was used as the agglomeration method for hierarchical clustering.

Gene ontology (GO) analysis for cluster 1 and 2 was conducted using ClueGO (v2.1.0) (5), a plug-in of Cytoscape (3.0.1) (42). When we searched for enrichment of GO terms, right-sided hypergeometric test was used with Benjamini-Hochberg correction on *P* value. The threshold of *P* value for biological process terms is 0.005 for cluster 1 and 0.0001 for cluster 2 and for cellular component and molecular function terms is 0.05 for both clusters. GO terms were restricted to levels 3 to 8. GO annotation evidence under Inferred from Electronic Annotation (IEA) was excluded. The restriction threshold for GO term connection (κ score), which determines the association strength between the terms, was set to 0.4. GO term fusion function and grouping function were

enabled to reduce term redundancy. The size of each term node represents its relative *P* value, and the color represents the group to which it belongs. The leading term of each GO term group was assigned based on the highest significance. GO, KEGG, and Interpro analysis of all changing transcripts was performed using the DAVID Bioinformatics Resource (v6.7) developed by the NIAID (NIH).

RESULTS

To identify novel regulators of heart failure, we carried out a quantitative proteomic analysis of cardiac chromatin. Adult male mice were subjected to transverse aortic constriction, and their cardiac function was monitored by echocardiography for a period of 2–6 wk. In this model, mice develop progressive cardiac disease that is hallmarked by two phases: a compensatory hypertrophy phase, in which the muscle increases in size while the function of the organ is preserved, and a failure phase, in which the left ventricular chamber size increases and function deteriorates. Using high-resolution mass spectrometry, we identified 176 proteins displaying altered association with chromatin as mice progressed into hypertrophy and failure (34). Unsupervised clustering of these proteins, following Z-score transformation of mass spectrometric intensities, revealed nine modules of proteins exhibiting distinct behaviors during pressure overload-induced heart failure. We performed GO analyses and literature searches on these proteins and identified the histone methyltransferase Smyd1 as a potential novel regulator of cardiac gene expression in the diseased heart. As detected in our mass spectrometry analysis, chromatin-bound Smyd1 is significantly increased during hypertrophy, and this increase is maintained in heart failure (Fig. 1*A*).

Subsequent Western blotting confirmed that both nucleoplasmic- and chromatin-bound populations of Smyd1 increased in the hypertrophic and failing mouse heart [Fig. 1, *B* and *D*; pressure overload hypertrophy is associated with manifold changes in histone variant expression, but these do not include a change in the total abundance of histone H3, as detected by Western blotting (11)]. To investigate whether changes in Smyd1 levels were conserved across species and hypertrophic stimulus, we treated isolated NRVMs with ISO, a β-adrenergic receptor agonist that induces pathological cell growth. Similar to what was observed in mice after pressure overload, ISO [at a dose shown to induce myocyte hypertrophy (31)] increased total cellular Smyd1 expression in isolated cells (Fig. 1, *C*, *top*, and *D*). Analysis of subcellular compartments revealed that this increase occurred preferentially in the nuclear fraction and was associated with Smyd1 translocation from cytosol to nucleus and/or resulted from a change in protein stability in one of these compartments (Fig. 1*C*, *middle* and *bottom*, and 1*D*) although the data did not allow us to distinguish between these possibilities.

To investigate the role of Smyd1 in the adult heart, we next took a genetic approach. Inducible, myocyte-specific *smyd1*-null mice were generated by first engineering a *smyd1* floxed mouse with the second and third exons targeted for deletion. These mice were bred to homozygosity, after which they grew and reproduced normally with no obvious phenotype. The *smyd1flox/flox* mice were then crossed with transgenic mice expressing the α-myosin heavy chain promoter-driven *Cre* gene flanked by two modified estrogen receptors. As described (43), this mouse expresses the resultant "MerCreMer" protein only in the heart. Upon treatment with the estrogen analog tamoxifen, this protein translocates to the nucleus and executes *loxp*-targeted

recombination activity. The experimental scheme for these mice is shown in Fig. 2*A* and confirmed by PCR in Fig. 2*B*. Resulting *MerCreMer-smyd1[flox/flox]* mice grew and reproduced normally in the absence of tamoxifen. Treatment of the mice with tamoxifen for 3 wk, followed by 1 wk on a normal diet, led to robust loss of Smyd1 at the protein level in the heart, while skeletal muscle Smyd1 was left unaffected (Fig. 2*C*).

Loss of Smyd1 in the adult heart induced fulminant heart failure, as proven by rapid deterioration of ejection fraction (Fig. 3*A*) and dilation of the left ventricle (Fig. 3*B*). Postmortem analyses revealed significant increases in heart weight:body weight ratio as early as 2 wk (the earliest time point examined) after halting tamoxifen treatment. This index of cardiac pathology had nearly doubled by the 9–10-wk time point (Fig. 3*C*) and was driven by a concomitant increase in cardiomyocyte size, as measured by wheat germ agglutinin staining (Fig. 3, *D* and *E*, *left*). Fibrosis also resulted in the wake of Smyd1 depletion (Fig. 3*E*, *right*). Because others have reported a transient cardiomyopathy in MerCreMer mice in the absence of a floxed allele (22), we monitored cardiac function in a cohort of these mice in our study, observing no heart failure and cardiac hypertrophy phenotypes (Fig. 3*F*) following an identical treatment protocol used to generate the Smyd1-deficient animals.

We attempted to carry out knockdown studies targeting Smyd1 in neonatal cardiac myocytes but were unsuccessful; doses of siRNA sufficient to knockdown Smyd1 induced significant cell death (data not shown). We speculate this finding reflects a differential role of the protein in neonatal development; indeed it is in agreement with the fact that the original Smyd1 germline knockout animals were lethal at the embryo stage (14). The progressive growth of the heart is also evident by H and

E staining of tissue sections at 9 wk (Fig. 3*G*) after tamoxifen administration. Heart rate was unaffected by loss of *smyd1* (data not shown). To investigate whether deletion of *smyd1* led to myocyte proliferation, mice were administered BrdU, which is incorporated into DNA during mitosis. We observed no BrdU-positive myocytes in *MerCreMer-smyd1*^{flox/flox} mice on normal or tamoxifen diet (Fig. 3*H*), suggesting that cellular proliferation was not a significant contributor to the increase in muscle mass observed following loss of *smyd1*. We have followed the *MerCreMer-smyd1*^{flox/flox} mice out to 10 wk after return to normal chow and observed no increase in mortality compared with either tamoxifen-treated *MerCreMer* mice with no floxed allele or untreated mice (that is, no mice died in any of these groups up until that time point). In separate experiments (data not shown), we observed that subjecting the Smyd1-depleted mice to pressure overload led to a more precipitous decline in cardiac function compared with either Smyd1 depletion or aortic banding alone although further investigation will be necessary to determine how the stress of Smyd1 loss interacts with the pathophysiological response to aortic banding.

Cardiac hypertrophy has been reported to involve a shift in gene expression wherein the diseased adult heart converts to a transcriptome reminiscent of the fetal myocardium (40). The mechanisms controlling this transition at the level of chromatin are unknown. Because Smyd1 was found to be upregulated in heart failure (Fig. 1), we reasoned that this protein may be an endogenous inhibitor of this disease-associated gene expression program. To test this hypothesis, we investigated expression of pathological genes in *smyd1*-null hearts. RT-PCR measurements of transcript abundance for α/β-myosin heavy chains, atrial natriuretic factor, sarcoplasmic reticulum

Ca^{2+} ATPase, glucose transporters (type 1 and 4), and carnitine palmitoyltransferase (type 1a and 1b) revealed significant changes following loss of *smyd1* (Fig. 4, *A–H*). With the exception of Glut1, these changes were progressive and mirrored the deterioration of cardiac function in these mice (Fig. 4). We next examined signaling molecules implicated in Smyd1 function, including the Smyd1 chaperone HSP90 [Fig. 4*I* (2, 48)], which was increased, and the Smyd1 interactor p53 [Fig. 4*J* (2, 21)], which was decreased, in the *smyd1*-null heart. Although ostensibly incapable of functional compensation at the cell and organ level, loss of *smyd1* led to a modest increase in expression from the closely related family member, Smyd2 [Fig. 4*K*; note that Smyd2 is dispensable for cardiac development and mature function (8), and pressure overload did not induce changes in its abundance in the wild-type heart as shown in Fig. 1*E*]. To investigate whether loss of *smyd1* affects global levels of histone posttranslational modifications, we performed Western blotting for well-characterized histone methylation events in total cell lysates. Interestingly, loss of *smyd1* had no effect on global levels of the heterochromatin mark, H4K20Me3 (Fig. 4*L*), the euchromatic mark, H3K4Me3 (Fig. 4*M*), or the constitutive heterochromatic mark, H3K9Me3 (Fig. 4*N*).

To further investigate the signaling pathways and specific gene expression events regulated by Smyd1, we took an unbiased transcriptome analysis approach. Hearts were harvested at an early (2 wk) and late (9 wk) time point after tamoxifen administration, and right and left ventricles were collected. These time points correspond to mild and severe cardiac dysfunction, respectively, following loss of *smyd1* in these animals. mRNA was purified, and expression microarrays were conducted on these samples, with the age-matched normal diet-fed animal serving as a

time- and chamber-specific control (4 animals were examined in each group). This dataset is a resource for understanding molecular targets of Smyd1 but also an important global measurement of the transcriptome remodeling that occurs during the transition of the heart into hypertrophy and on to failure.

Before focusing on individual genes, however, we wanted to determine the global role of Smyd1 to influence gene expression. To do this, all transcriptomic data (including early and late time points, right and left ventricles) were plotting as a heat map and clustered in an unsupervised manner. The result (Fig. 5*A*) displays upregulated (red areas, so called cluster 2) and downregulated (green areas, cluster 1) transcripts. We next performed GO and KEGG analysis on these two major clusters (Fig. 5, *B* and *C*) to generate network maps of key processes regulated by Smyd1. Network nodes are sized according to *P* value and colored according to the most significantly enriched term in each group. There are two salient observations: first, downregulated transcripts are enriched for genes involved in metabolic processes (Fig. 5*B*); second, upregulated genes in the knockout are enriched for developmental and muscle growth genes (Fig. 5*C*), providing unbiased evidence in support of the conclusion that Smyd1 is a global repressor of development and growth in the adult heart.

We next examined each of the microarray datasets individually using GO and KEGG (followed by Interpro domain analyses for the implicated pathways), revealing the pathways implicated in a chamber-specific manner over time, as loss of *smyd1* induces cardiac hypertrophy and failure. Bioinformatic analyses revealed a significant enrichment in molecules involved in transcriptional repression, muscle development, growth, TGF-β signaling, and extracellular matrix (Fig. 5, *D* and *E*),

supporting our histological observation of increased collagen deposition in the *smyd1*-depleted myocardium (Fig. 3*E*). Indeed, RT-PCR was used to confirm changes in expression of several of the genes measured by microarray, including three members of the TGF-β family (Fig. 5*F*), further validating the role of Smyd1 to repress growth signaling and gene expression in the normal heart.

To identify gene targets directly regulated by Smyd1 and to further investigate the mechanism by which Smyd1 controls gene expression, we examined the *tgfbeta3* and *nppa* loci using ChIP-PCR. Previous studies had shown that sequences within 1 kb 5′ of the transcriptional start sites of *tgfbeta3* (28) and *nppa* (10) were essential for transcription. Consistent with these observations, we observed significant enrichment of Smyd1 binding upstream of both the *tgfbeta3* and *nppa* transcriptional start sites (Fig. 5, *G* and *H*, *top*), demonstrating direct recruitment of Smyd1 to their promoter regions. It should be noted that we also performed ChIP further 5′ of these regions because previous work had identified critical control regions substantially upstream of the transcriptional start site for *tgfbeta3* (28) and particularly for *nppa* (19). Contrary to our initial hypothesis, Smyd1 binding did not correlate with enrichment of histone H3 lysine 4 trimethylation across the region (Fig. 5, *G* and *H*, *bottom*). However, luciferase analyses confirmed that Smyd1 binding to these regions repressed transcription of *tgfbeta3* and *nppa* (Fig. 5*I*). To determine whether Smyd1 modulates these genes in the heart, we investigated localization of Smyd1, as well as multiple histone marks, to the transcription start site region in *tgfbeta3* and *nppa* using ChIP-PCR. Loss of Smyd1 was also associated with a decrease in multiple histone marks, concomitant with a decrease in total histone H3, in

the promoters of these genes (data not shown), which we interpret as evidence of general nucleosome depletion in the promoter regions, possibly as part of transcriptional activation, although future studies will be required to test this conjecture. Last, ChIP-PCR against genes not targeted by Smyd1 [using primers (24, 29) validated to target regulatory regions of these genes from previous publications] shows a lack of enrichment at beta-actin and beta-tubulin, supporting specificity in targeting.

Because the absence of Symd1 induced rapid heart failure in the adult mouse, we last sought to determine whether overexpression of Smyd1 would be sufficient to prevent hypertrophy. There are two isoforms of Smyd1 in the mouse heart; Smyd1a differs from Smyd1b by the presence of a 13-amino-acid region in the E-T domain of the former, a feature that is conserved in the single isoform of Smyd1 in humans (Fig. 6*A*). Adenoviral-mediated overexpression of Smyd1a or Smyd1b was successful (at varying MOI) to increase Smyd1 protein levels (Fig. 6, *B–D*). We detected no effect of this overexpression on basal cell phenotype in isolated cardiac myocytes. However, overexpression of Smyd1a (Fig. 6, *E* and *F*), but not Smyd1b (Fig. 6, *G* and *H*), was sufficient to block PE-induced hypertrophy. Overexpression of Smyd1a or Smyd1b also altered the expression of some genes often associated with cardiac pathology (Fig. 6*I*). For example, basal levels of SERCA and α/β-myosin heavy chain (MHC) were downregulated (empty virus infection blocked the PE-induced upregulation of β-MHC we observe in untransfected cells), and the increase in ANF induced by PE was blocked. In agreement with the actions of TGF-β3 as a potential downstream target of Smyd1, the basal expression of this gene (along with TGF-β1) was diminished after Smyd1 activation. These findings indicate that, although only Smyd1a was effective to prevent

hypertrophy in isolated cardiac myocytes, both isoforms can regulate gene expression, and their actions impact both basal and agonist-induced transcript abundance.

DISCUSSION

During normal development, a series of epigenetic transitions converts multipotent progenitor cells into the various lineages of the heart, including cardiac myocytes. Soon after birth in mammals, cardiomyocytes become binucleated and stop dividing. From this point forward, their response to injury or stress is (almost always) to either die or increase in size. Although the latter can be beneficial to maintain the function of the organ, cellular hypertrophy—and the concomitant conversion to a pathologic (and perhaps more developmentally primitive) transcriptome and proteome that accompanies it—ultimately leads to heart failure. In the present study, we provide in vivo evidence that addresses key elements of this mystery of cardiac function. How does the healthy adult heart keep developmental genes silenced? Our data demonstrate that loss of Smyd1 in the adult heart removes a repression on such genes, leading to extensive transcriptome remodeling and precipitous heart failure. In the normal heart, Smyd1 is upregulated during the hypertrophic phase of disease, probably as part of an endogenous defense mechanism that is ultimately overwhelmed when the heart transitions to failure.

Previous investigations have shown unequivocally that chromatin modification can induce hypertrophic growth, as well as restrict it, including in physiological and pathological conditions (3, 4, 15, 23, 36, 50). One study showed that the histone methyltransferase JMJD2A is a positive regulator of hypertrophy, with overexpression

57

leading to exacerbated response to hypertrophy and knockout protecting against it (53). However, this protein is not cardiac specific, likely explaining the lack of a phenotype in these JMJD2A transgenic or knockout animals at baseline. Despite this exciting finding, until now no study to our knowledge has demonstrated a muscle-specific chromatin protein that restricts adult cardiac growth. Our results are a step toward reconciling cardiac-specific gene expression with the cell type-independent, genome-wide actions of many chromatin modifiers. How do cardiac transcription factors, active in the adult heart, avoid spurious transcription of developmentally silenced genes? Our data demonstrate that this occurs partly through the actions of Smyd1, which we observe to be an endogenous inhibitor of a panel of pathological and developmental genes in the adult myocardium.

There are five Smyd family members expressed in mammals, but only Smyd1 is restricted to striated muscle. The human *smyd1* gene produces a single transcript, which shares this striated muscle specificity with two of the three murine *smyd1* transcripts, Smyd1a and b (the third, murine Smyd1c, being restricted to T cells). Previous studies have shown Smyd1 to be necessary for early heart morphogenesis (14) and to be transcriptionally regulated by serum response factor and myogenin (26). Other investigations have examined Smyd isoforms in striated muscle, demonstrating a role for zebrafish Smyd1b in skeletal and cardiac myofibril assembly (27, 46). In our study, Smyd1b in the mouse had no effect on hypertrophic cell growth (Fig. 6, *G* and *H*), whereas the Smyd1a isoform prevented pathological changes in myocytes exposed to PE (Fig. 6, *E* and *F*). In addition to these novel phenotypic findings regarding the ability of Smyd1 to inhibit disease progression, this study

supports a paradigm shift in the basic biological function of Smyd1. Smyd family members were originally thought to activate gene transcription through methylation of histone H3 lysine K4; however, to date this enzymatic activity has only been reported to occur in vitro. Conversely, the notion of Smyd1 as a transcriptional repressor was first suggested by Gottlieb et al. (14), without direct in vivo evidence for this action in adult myocardium. The present study demonstrates the in vivo activity of Smyd1 as a transcriptional repressor of a host of genes (Fig. 5, *A–C*). Furthermore, we show by ChIP and luciferase reporter assay that Smyd1 binds directly to the promoter regions of *tgfβ3* and *nppa*/ANF and inhibits their expression. These findings were not associated with significant changes in H3K4Me3 globally or at targeted loci we evaluated, calling into question the model in which Smyd1 targets this histone modification in vivo.

Disease models in vivo and in cell culture have complementary strengths and weaknesses for analyzing gene expression. In Smyd1-deficient hearts, we observed a progressive reprogramming of fetal genes, with β-MHC and ANF increased and α-MHC and SERCA2A decreased. These changes were consistent with the expected trends for these genes during heart failure. However, unlike α-MHC and ANF, neither β-MHC nor SERCA2A responded as the anticipated trend in the Smyd1 overexpression scenario. There are several possible explanations for these observations. For example, the in vivo situation may result from, not only the direct actions of Smyd1, but also indirect actions that require more time to develop to the observed endpoint of heart failure. Also, the overexpression experiment that led to elevated levels of Smyd1 may not be optimal in that we are attempting to modify baseline gene expression in addition to agonist-induced changes. Future studies using transgenic and/or viral strategies to modulate

Smyd1 levels in the heart may be required to fully address the regulation of these specific gene targets.

Because of the known role of Smyd proteins to regulate sarcomere function in vivo (48), we sought to examine nonnuclear populations of Smyd1 during disease. Our findings indicate that changes in nuclear Smyd1 in the setting of heart failure involve a concomitant decrease in cytosolic Smyd1 (which under basal conditions is comparable to nuclear Smyd1); this occurs along with an increase in total cellular Smyd1 protein levels. Furthermore, and also relevant to the actions of Smyd1 outside of the nucleus, loss of Smyd1 led to increased extracellular fibrosis, which could be in part due to the actions of Smyd1 at the membrane, in addition to transcriptional actions of this protein (Fig. 7). In contrast, overexpression experiments failed to alter the general myofibrillar architecture while being sufficient to block agonist-induced hypertrophy. Therefore, further work is required to distinguish genomic from nongenomic actions of Smyd1, given the convincing data from previous studies implicating this family of proteins in cytoplasmic and membrane, i.e., nonhistone, targets. Indeed on the basis of our work with histone posttranslational modifications in this study, we think it highly likely that many of the targets of Smyd1 are unidentified residues on histones and/or unknown proteins in the nucleus and beyond.

It is interesting to speculate on the binding partners with which Smyd1 interacts to achieve specificity in its intracellular tasks. Previous investigations demonstrated roles for p53 and Hsp90 in Smyd family function (20, 48), in addition to being independently implicated in cardiac hypertrophy (25, 41), and indeed these proteins underwent altered expression following Smyd1 depletion in our study. Furthermore, we

directly tested the hypothesis that global histone H3K4 trimethylation levels would be altered in the knockout mice; they were not, contrary to our expectation. We did, however, see alteration in levels of H3K4Me3 around key cardiac genes *TGFβ* and *ANF* following Smyd1 overexpression, suggesting that there is focal regulation of histone marks. In our view, the most likely explanation for these observations is that global levels of chromatin accessibility are modulated by a cadre of proteins, rather than a single enzyme or pathway, and, when Smyd1 is depleted, other chromatin modifiers are engaged to maintain a modicum of control on gene expression. Our observation that transcriptome remodeling after Smyd1 depletion leads to massive changes in cardiac gene expression, but not wholesale transcriptional chaos (for instance, permanently silenced noncardiac genes are not activated en masse), supports a buffering capacity in the epigenomic machinery that preserves the cardiac phenotype.

A recent emergence (16, 17, 38, 49) of studies examining chromatin marks and transcription factors in cell systems and using recombinant proteins notwithstanding, genome-wide positioning of cardiac transcription factors in vivo, using ChIP-seq approaches, have been lacking. Indeed we were unsuccessful in our attempts to optimize immunoprecipitation of endogenous Smyd1 to sufficient levels to enable DNA sequencing, probably attributable to limitations of the available reagents and abundance and localization of endogenous Smyd1 (as discussed elsewhere, the protein partitions between nucleus and cytosol in cardiomyocytes).

Studies from human patients with heart failure indicate that significant transcriptome and phenotypic remodeling is possible even in the sick heart (30) through interventions such as ventricular assist devices (7, 51), which mechanically unload the

ventricle. Indeed mechanical unloading of the heart in humans has been shown to return endogenous Smyd1 expression to basal levels following its upregulation in the diseased human heart (6). These observations provide rationale for strategies to globally remodel chromatin, but a limitation has been that this remodeling must be done in a cell type-specific manner, to avoid unwanted side effects and to improve efficacy. The identification of a muscle-specific histone modifier that represses endogenous cardiac growth provides an exciting molecular target to overcome this limitation. Recent studies have indicated that Smyd1 plays an important role in other types of striated muscle; its absence during development (Myf5-driven) undermined myoblast differentiation (35), whereas disruption in the committed myoblast lineage (Myf6-driven) led to fast-twitch skeletal muscle weakness, hypotrophy (in contrast to the hypertrophy in the heart), and myofibrillar disarray, among other subcellular abnormalities (45). Nevertheless, given the strong sequence conservation between mouse and human Smyd1 (94.49% similarity at the amino acid level; Fig. 6*A*) and the fact that Smyd1 is upregulated in human heart failure (1, 6), it is an exciting indication that this pathway may be operative in humans. The present study examines the molecular basis for how Smyd1 functions in vivo and provides compelling evidence that it may serve as a novel, muscle-specific therapeutic target.

GRANTS

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

S.F., H.O.T., and T.M.V. conception and design of research; S.F., T.K., T.L.R., M.R.G., T.T., M.R.M., R.G., S.J., and S.R. performed experiments; S.F., T.K., M.R.G., H.C., Y.W., and T.M.V. analyzed data; S.F., T.K., M.R.G., H.C., Y.W., H.O.T., and T.M.V. interpreted results of experiments; S.F., M.R.G., H.C., and T.M.V. prepared figures; S.F. and T.M.V. drafted manuscript; S.F. and T.M.V. edited and revised manuscript; S.F., T.K., T.L.R., M.R.G., H.C., T.T., M.R.M., R.G., S.J., S.R., Y.W., H.O.T., and T.M.V. approved final version of manuscript.

Figure Legends

Fig. 1. Identification of Smyd1 as a novel participant in pathological cardiac hypertrophy and failure. *A*: among hundreds of chromatin-bound proteins quantified during pressure overload-induced hypertrophy and heart failure, Smyd1 abundance increased during compensatory hypertrophy and failure, as measured by mass spectrometry-based peptide abundance from cardiac chromatin. This trend was confirmed by Western blotting (*B*); we also observed an increase in nucleoplasmic (i.e., nuclear localized, not bound to chromatin) Smyd1 during heart failure. IB, immunoblot. *C*: to determine whether activation of Smyd1 was conserved across species and stimulus, we examined expression and localization following isoproterenol (ISO; 1 μM for times indicated) treatment of ventricular myocytes. *D*: quantitation of Western blots from *B* and *C*. *E*: Smyd2 expression is unchanged in the setting of pressure overload-induced heart failure; $n = 4–6$ per group for all Westerns; bars are SE, $*P < 0.05$.

Fig. 2. Generation of cardiac-specific, tamoxifen-inducible *smyd1*-null mice. *A*: positional cloning was used to insert *loxp* sites flanking *exons 2* and *3* of *smyd1* to produce the target allele. These mice were bred to homozygosity and then crossed with transgenic mice expressing the Cre recombinase positioned between two copies of the tamoxifen-responsive modified estrogen receptor and driven by the α-myosin heavy chain (MHC) promoter (so called Mer-Cre-Mer mice). *B*: PCR genotyping confirmed genetic manipulation (location of primers are indicated in *A*). Resulting mice express Mer-Cre-Mer in adult cardiomyocytes, and, upon treatment with tamoxifen (Tmx), this protein excises the floxed region of the mutant *smyd1* allele. Mice fed a tamoxifen diet for 3 wk

followed by 1 wk on a normal diet demonstrate near complete loss of Smyd1 protein specifically in the heart; Smyd1 in skeletal muscle is unaffected (*C*). Diagonal lines indicate removal of gel lanes. WT, wild-type.

Fig. 3. Loss of Smyd1 in the adult heart induces progressive hypertrophy and failure at the organ and cell level. Cardiac function in *smyd1*$^{flox/flox}$*Cre*$^{+/-}$ mice was measured by echocardiography (ECHO). A baseline reading was taken in each mouse before 3 wk on tamoxifen or regular diet. After 1 wk back on regular diet in the experimental group, ECHO measurements were made on a weekly basis in both groups. We observed a progressive decline in left ventricular ejection fraction (*A*) coupled with a progressive increase in left ventricular end-diastolic dimension (LVEDd) (*B*) as a result of *smyd1* deletion (*n* = 37 control mice and 36 tamoxifen=treated mice for *A* and *B*; *P* < 0.05 vs. regular diet at same time point). Loss of Smyd1 also induced an increase in muscle mass at the whole organ (as measured by heart weight:body weight ratio, HW/BW, *P* < 1E-5; *C*) and individual cardiac cell level as measured following wheat germ agglutinin staining (*D* and *E*, *left*, 10 wk after tmx; bar = 50 μm), concomitant with an increase in fibrotic deposition as measured by trichrome staining (*E*, *right*, 8 wk after tmx; bar = 100 μm); *n* values for HW/BW and cell size are indicated, the former indicating number of animals and the latter indicating number of cells (from a total of 4 animals per group). KO, knockout. *F*: *MerCreMer-smyd1*$^{wt/wt}$ mice (i.e., mice with no floxed alleles) develop neither cardiac dysfunction nor hypertrophy when administered the same tamoxifen protocol. *G*: marked chamber dilation is observed in Smyd1-deficient mice at 8–10 wk after return to normal chow. *H*: *MerCreMer-smyd1*$^{flox/flox}$ mice were injected with 5-

bromo-1-(2-deoxy-β-d-ribofuranosyl) (BrdU) to label actively replicating cells. Small intestine positive control shows strong BrdU incorporation. We observe incorporation of BrdU in neither the hearts of *MerCreMer-smyd1^flox/flox* mice fed a normal diet (*middle*) nor those with the same genotype on a tamoxifen diet for 10 wk (*bottom*; bar = 50 µm).

Fig. 4. Smyd1 is a previously unknown regulator of a conserved transcriptional program underlying heart failure. *A–H*: fetal genes were measured by RT-PCR and values expressed as fold change relative to mice not fed tamoxifen. *$P < 0.05$; $n = 5–8$ for no tmx, $n = 6$ for post tmx 2 wk, and $n = 7$ for post tmx 9–10 wk. ANF, atrial natriuretic factor. *I–N*: as determined by Western blotting, loss of Smyd1 increased expression of Hsp90, decreased expression of p53, resulted in an upregulation of Smyd2, and had no effect on global levels of histone H4 (K20) and H3 (K4 and K9) trimethylation. *O*: quantitation of *I–N*; *$P = 0.002$.

Fig. 5. Transcriptome analyses reveal Smyd1 to be a transcriptional repressor of a core set of developmental genes. Transcriptome analyses were performed separately on right ventricles (RV) and left ventricles (LV) from mice with deletion of Smyd1 at 2 wk and 9 wk after removal from tamoxifen diet (controls were normal diet-fed littermates). *A*: all genes measured in both groups, RV and LV, are displayed in heat map format (red is upregulation, green downregulation, and black statistically unchanged) and clustered according to similar behavior. 2 clusters were defined as indicated, and the functionality of genes in those clusters was determined by gene ontology (GO) analysis. The networks for cluster 1 (*B*) and 2 (*C*) for the biological process ontology are shown. In the

network figures, node size indicates the *P* value, and linkage between two terms indicates the relatedness, as determined by κ statistics. The color indicates functional groups with each group represented by their most significant leading term. All terms shown in the network image of cluster 1 are filtered at $P < 0.005$; those in cluster 2 are filtered at $P < 0.0001$. Bioinformatic analyses of transcripts with altered expression in the LV, at *week 2* (*D*) and *week 9* (*E*) after Tmx treatment, show significant enrichment in genes involved in extracellular matrix remodeling, fibrosis, and transcriptional repression. MF, molecular function; BP = biological process; CC, cellular component; KEGG, KEGG analysis; ITP, Interpro analysis. Microarray results for several of these transcripts were subsequently validated by RT-PCR (*F*: all those we attempted to validate are shown; $n = 3–6$/group; *$P < 0.05$ vs no tamoxifen group). Chromatin immunoprecipitation (ChIP) and qPCR for Smyd1 (using FLAG antibody) show enrichment in the promoter region of target genes *tgfbeta3* (*G*) and *nppa* (*H*); however, no corresponding enrichment of histone H3 lysine K4 trimethylation was detected in these regions (*bottom*). *$P ≤ 0.05$. *I*: luciferase reporter assay using the *tgfbeta3* and *nppa* promoters confirms that Smyd1 acts as a transcriptional repressor by inhibiting transcription of these genes; $n = 6$/group; *$P < 0.05$. *J*: as a negative control, Smyd1 is enriched by ChIP-PCR at neither β-tubulin nor β-actin using primers shown previously to target the regulatory regions upstream of these genes [−3 kb for β-tubulin (24), −73 bp for β-actin (29)].

Fig. 6. Overexpression of Smyd1 attenuates cellular hypertrophy in vivo. *A*: there are 2 splice variants of Smyd1 in mouse heart, differing by a 13-amino-acid sequence present

in Smyd1a and absent in Smyd1b, but conserved in the single human Smyd1 transcript. *B–D*: adenoviral infection of myocytes led to robust Smyd1 expression, as confirmed by RT-PCR and Western blotting. MOI, multiplicity of infection. *E* and *F*: overexpression of Smyd1a had no effect on basal cell size but prevented phenylephrine (PE)-induced myocyte hypertrophy (phalloidin staining; scale bar = 100 μm; quantified from ~100 cells per group in 3 independent experiments; *P < 0.001). *G* and *H*: overexpression of Smyd1b required higher MOI and had no effect on PE-induced cell growth (quantified as described in *F*). *I*: RT-PCR analysis of fetal genes and Smyd1 targets in neonatal rat ventricular myocytes after adenovirus infection in the presence and absence of PE. *$P ≤$ 0.05 compared with empty virus control without PE.

Fig. 7. Model for functions of Smyd1 in the adult myocardium. Under basal conditions, Smyd1 functions as a transcriptional repressor to inhibit cell growth and maintain cardiomyocyte size. Loss of Smyd1 in the adult myocardium leads to prohypertrophic signaling resulting in myocyte growth, fibrosis, and functional decline.
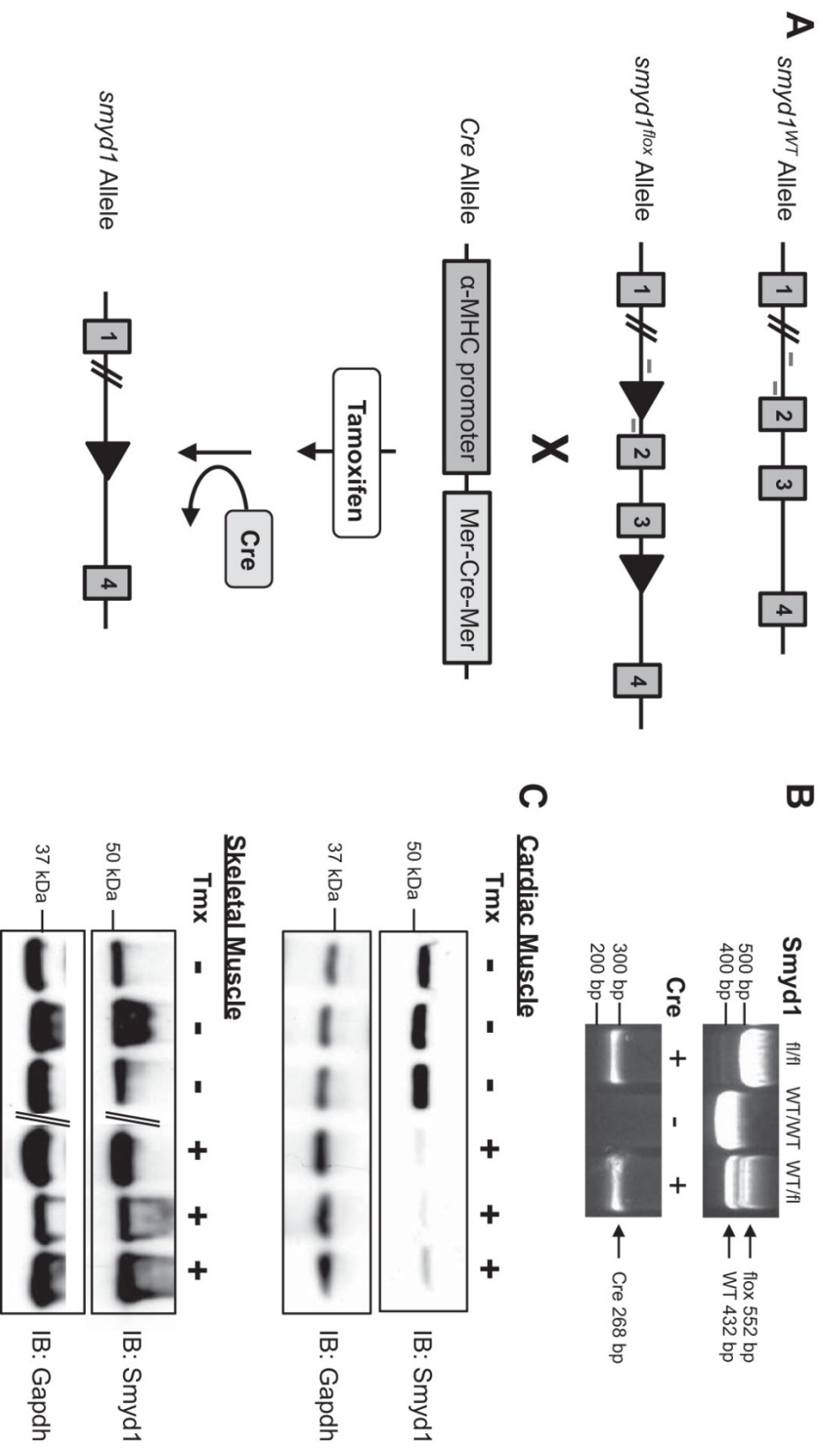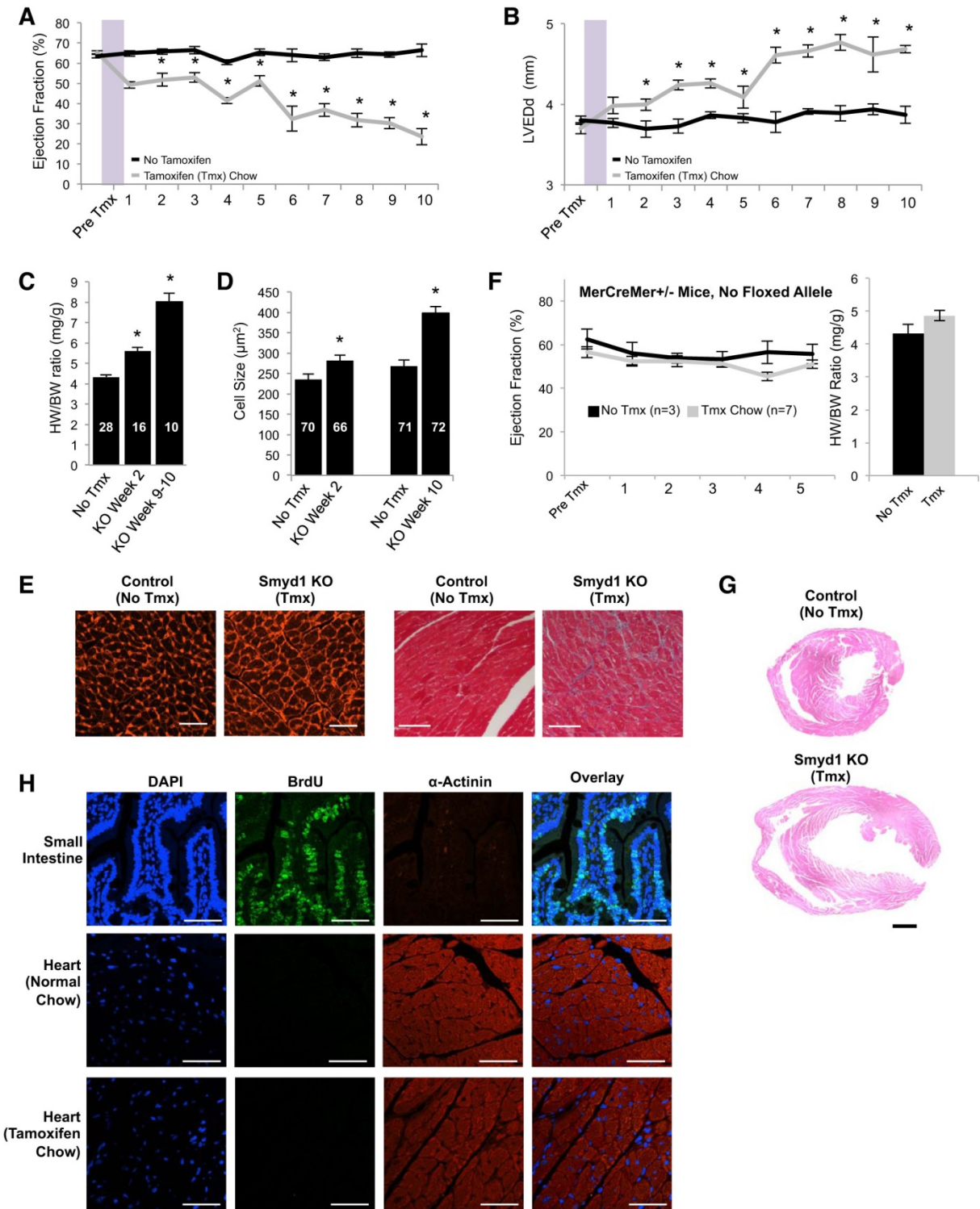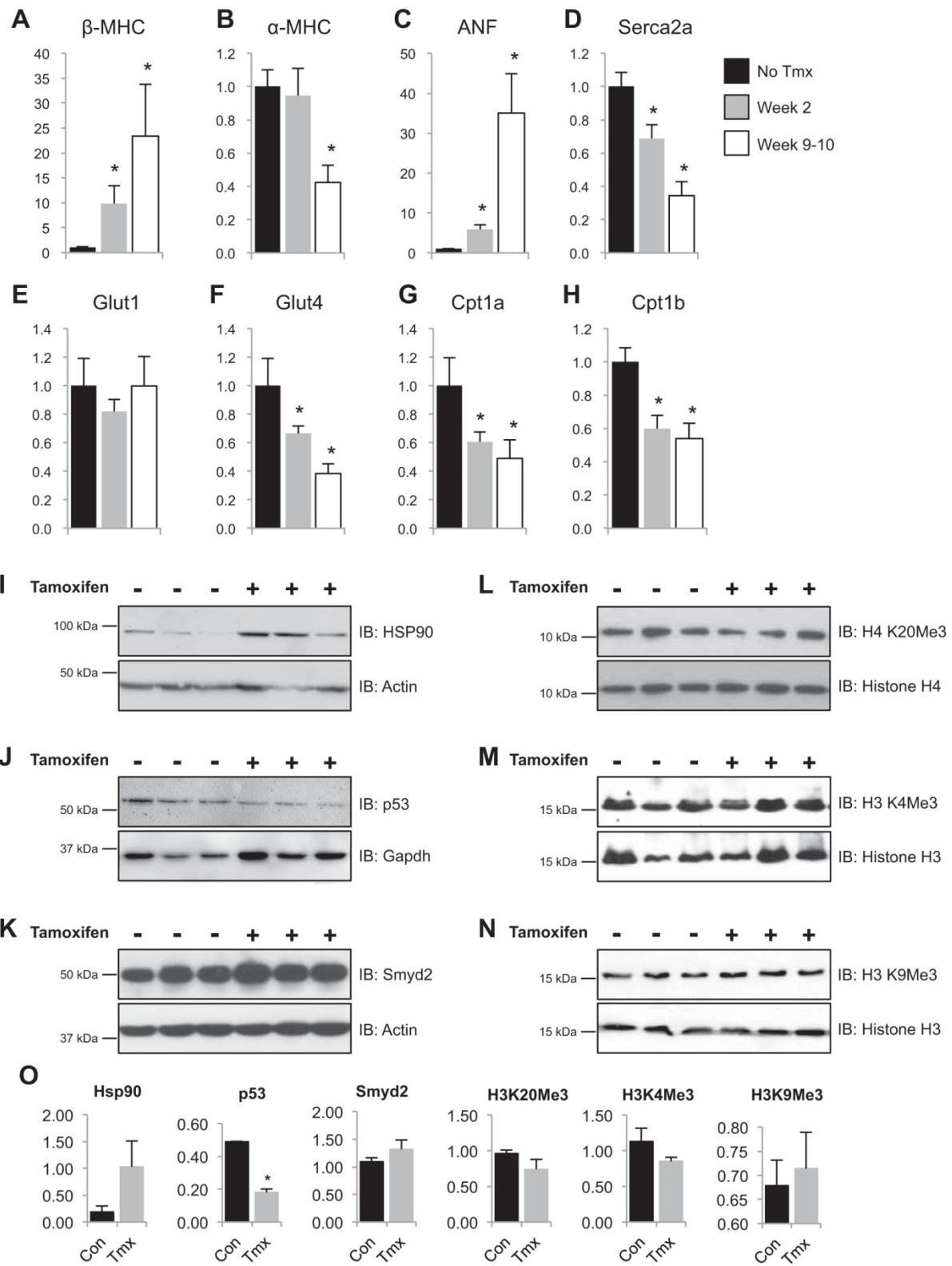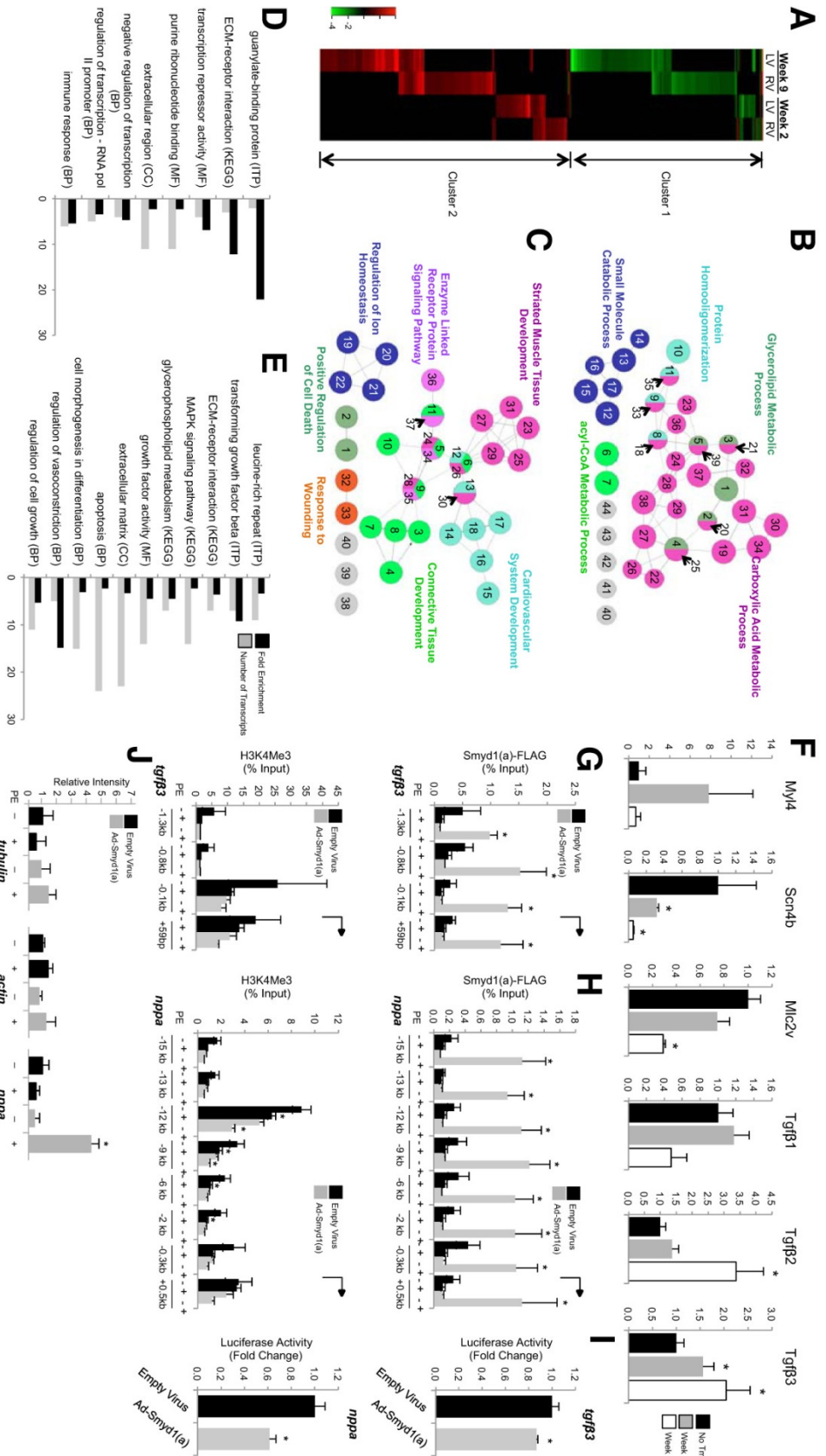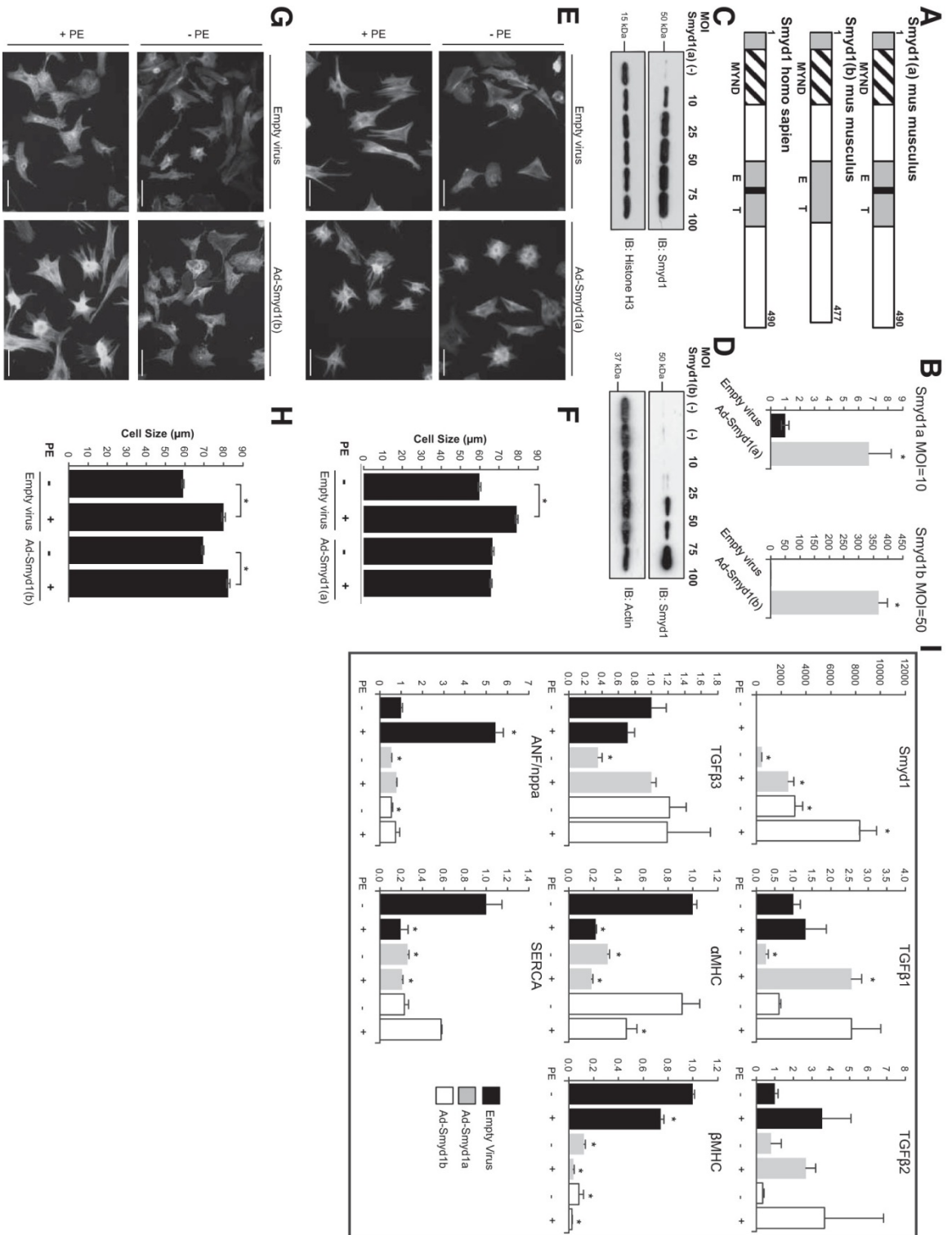
Fig. 1

69

Fig. 2

70
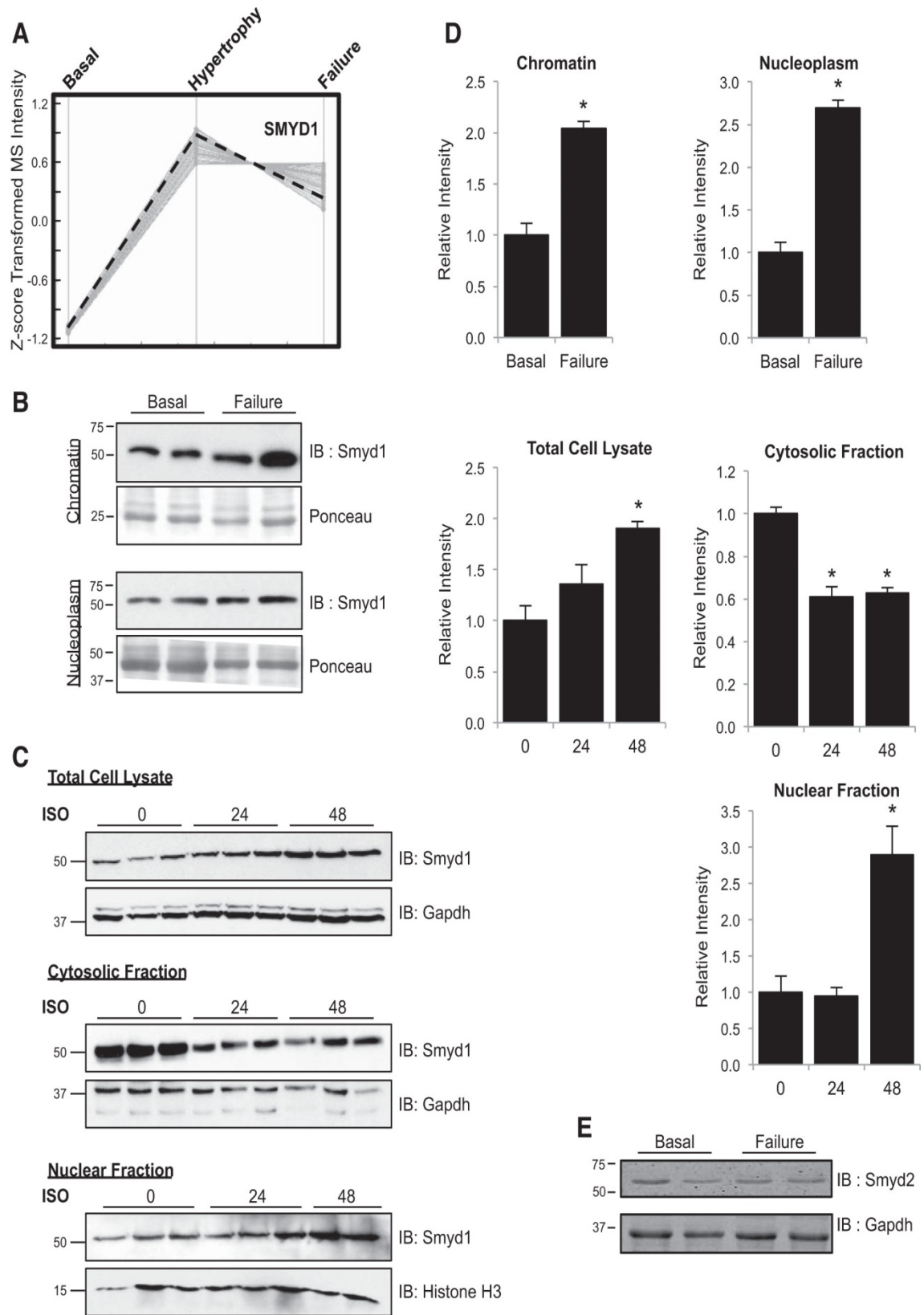
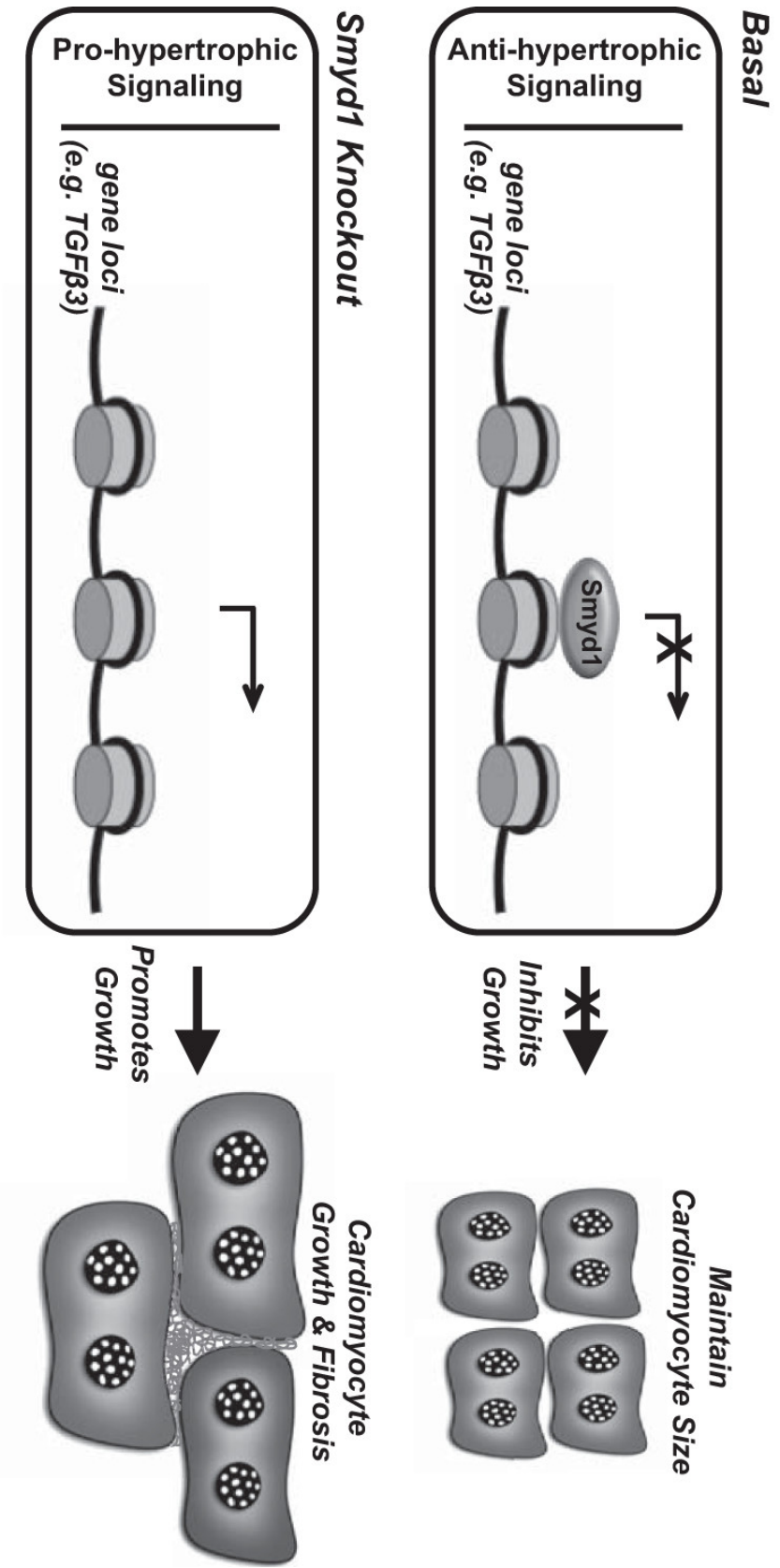Fig. 3

Fig. 4

Fig. 5

Fig. 6

Fig. 7

References

1. Abaci N, Gulec C, Bayrak F, Komurcu Bayrak E, Kahveci G, Erginel Unaltuna N. The variations of BOP gene in hypertrophic cardiomyopathy. *Anadolu Kardiyol Derg* 10: 303-309, 2010.

2. Abu-Farha M, Lambert JP, Al-Madhoun AS, Elisma F, Skerjanc IS, Figeys D. The tale of two domains: Proteomics and genomics analysis of SMYD2, a new histone methyltransferase. *Mol Cell Proteomics* 7: 560-572, 2008.

3. Ago T, Liu T, Zhai P, Chen W, Li H, Molkentin JD, Vatner SF, Sadoshima J. A redox-dependent pathway for regulating class II HDACs and cardiac hypertrophy. *Cell* 133: 978-993, 2008.

4. Anand P, Brown JD, Lin CY, Qi J, Zhang R, Artero PC, Alaiti MA, Bullard J, Alazem K, Margulies KB, Cappola TP, Lemieux M, Plutzky J, Bradner JE, Haldar SM. BET bromodomains mediate transcriptional pause release in heart failure. *Cell* 154: 569-582, 2013.

5. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pages F, Trajanoski Z, Galon J. ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25: 1091-1093, 2009.

6. Borlak J, Thum T. Hallmarks of ion channel gene expression in end-stage heart failure. *FASEB J* 17: 1592-1608, 2003.

7. Burkhoff D, Klotz S, Mancini DM. LVAD-induced reverse remodeling: Basic and clinical implications for myocardial recovery. *J Card Fail* 12: 227-239, 2006.

8. Diehl F, Brown MA, van Amerongen MJ, Novoyatleva T, Wietelmann A, Harriss J, Ferrazzi F, Bottger T, Harvey RP, Tucker PW, Engel FB. Cardiac deletion of Smyd2 is dispensable for mouse heart development. *PLoS One* 5: e9748-2010.

9. Dorn GW 2nd, Force T. Protein kinase cascades in the regulation of cardiac hypertrophy. *J Clin Invest* 115: 527-537, 2005.

10. Durocher D, Nemer M. Combinatorial interactions regulating cardiac transcription. *Dev Genet* 22: 250-262, 1998.

11. Franklin S, Chen H, Mitchell-Jordan S, Ren S, Wang Y, Vondriska TM. Quantitative analysis of the chromatin proteome in disease reveals remodeling principles and identifies high mobility group protein B2 as a regulator of hypertrophic growth. *Mol Cell Proteomics* 11: M111.014258 11: 2012.

12. Franklin S, Zhang MJ, Chen H, Paulsson AK, Mitchell-Jordan SA, Li Y, Ping P, Vondriska TM. Specialized compartments of cardiac nuclei exhibit distinct proteomic anatomy. *Mol Cell Proteomics* 10: M110.00070310: 2011.

13. Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Blaha MJ, Dai S, Ford ES, Fox CS, Franco S, Fullerton HJ, Gillespie C, Hailpern SM, Heit JA, Howard VJ, Huffman MD, Judd SE, Kissela BM, Kittner SJ, Lackland DT, Lichtman JH, Lisabeth LD, Mackey RH, Magid DJ, Marcus GM, Marelli A, Matchar DB, McGuire DK, Mohler ER 3rd, Moy CS, Mussolino ME, Neumar RW, Nichol G, Pandey DK, Paynter NP, Reeves MJ, Sorlie PD, Stein J, Towfighi A, Turan TN, Virani SS, Wong ND, Woo D, Turner MB. Heart disease and stroke statistics–2014 update: A report from the American Heart Association. *Circulation 129*: e28-e292, 2014.

14. Gottlieb PD, Pierce SA, Sims RJ, Yamagishi H, Weihe EK, Harriss JV, Maika SD, Kuziel WA, King HL, Olson EN, Nakagawa O, Srivastava D. Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis. *Nat Genet* 31: 25-32, 2002.

15. Hang CT, Yang J, Han P, Cheng HL, Shang C, Ashley E, Zhou B, Chang CP. Chromatin regulation by Brg1 underlies heart muscle development and disease. *Nature* 466: 62-67, 2010.

16. He A, Gu F, Hu Y, Ma Q, Ye LY, Akiyama JA, Visel A, Pennacchio LA, Pu WT. Dynamic GATA4 enhancers shape the chromatin landscape central to heart development and disease. *Nat Commun* 5: 4907-2014.

17. He A, Kong SW, Ma Q, Pu WT. Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. *Proc Natl Acad Sci USA* 108: 5632-5637, 2011.

18. Hill JA, Olson EN. Cardiac plasticity. *N Engl J Med* 358: 1370-1380, 2008.

19. Houweling AC, van Borren MM, Moorman AF, Christoffels VM. Expression and regulation of the atrial natriuretic factor encoding gene Nppa during development and disease. *Cardiovasc Res* 67: 583-593, 2005.

20. Huang J, Perez-Burgos L, Placek BJ, Sengupta R, Richter M, Dorsey JA, Kubicek S, Opravil S, Jenuwein T, Berger SL. Repression of p53 activity by Smyd2-mediated methylation. *Nature* 444: 629-632, 2006.

21. Jiang Y, Sirinupong N, Brunzelle J, Yang Z. Crystal structures of histone and p53 methyltransferase SmyD2 reveal a conformational flexibility of the autoinhibitory C-terminal domain. *PLoS One* 6: e21640-2011.

22. Koitabashi N, Bedja D, Zaiman AL, Pinto YM, Zhang M, Gabrielson KL, Takimoto E, Kass DA. Avoidance of transient cardiomyopathy in cardiomyocyte-targeted tamoxifen-induced MerCreMer gene deletion models. *Circ Res* 105: 12-15, 2009.

23. Kook H, Lepore JJ, Gitler AD, Lu MM, Wing-Man Yung W, Mackay J, Zhou R, Ferrari V, Gruber P, Epstein JA. Cardiac hypertrophy and histone deacetylase-dependent transcriptional repression mediated by the atypical homeodomain protein. *Hop J Clin Invest* 112: 863-871, 2003.

24. Kumar A, Choi KH, Renthal W, Tsankova NM, Theobald DE, Truong HT, Russo SJ, Laplant Q, Sasaki TS, Whistler KN, Neve RL, Self DW, Nestler EJ. Chromatin remodeling is a key mechanism underlying cocaine-induced plasticity in striatum. *Neuron* 48: 303-314, 2005.

25. Lee KH, Jang Y, Chung JH. Heat shock protein 90 regulates IkappaB kinase complex and NF-kappaB activation in angiotensin II-induced cardiac cell hypertrophy. *Exp Mol Med* 42: 703-711, 2010.

26. Li D, Niu Z, Yu W, Qian Y, Wang Q, Li Q, Yi Z, Luo J, Wu X, Wang Y, Schwartz RJ, Liu M. SMYD1, the myogenic activator, is a direct target of serum response factor and myogenin. *Nucleic Acids Res* 37: 7059-7071, 2009.

27. Li H, Zhong Y, Wang Z, Gao J, Xu J, Chu W, Zhang J, Fang S, Du SJ. Smyd1b is required for skeletal and cardiac muscle function in zebrafish. *Mol Biol Cell* 24: 3511-3521, 2013.

28. Liu G, Ding W, Neiman J, Mulder KM. Requirement of Smad3 and CREB-1 in mediating transforming growth factor-beta (TGF beta) induction of TGF beta 3 secretion. *J Biol Chem* 281: 29479-29490, 2006.

29. Maeda T, Hobbs RM, Merghoub T, Guernah I, Zelent A, Cordon-Cardo C, Teruya-Feldstein J, Pandolfi PP. Role of the proto-oncogene Pokemon in cellular transformation and ARF repression. *Nature* 433: 278-285, 2005.

30. Margulies KB, Matiwala S, Cornejo C, Olsen H, Craven WA, Bednarik D. Mixed messages: Transcription patterns in failing and recovering human myocardium. *Circ Res* 96: 592-599, 2005.

31. Mitchell-Jordan S, Chen H, Franklin S, Stefani E, Bentolila LA, Vondriska TM. Features of endogenous cardiomyocyte chromatin revealed by super-resolution STED microscopy. *J Mol Cell Cardiol* 53: 552-558, 2012.

32. Mitchell-Jordan SA, Holopainen T, Ren S, Wang S, Warburton S, Zhang MJ, Alitalo K, Wang Y, Vondriska TM. Loss of Bmx nonreceptor tyrosine kinase prevents pressure overload-induced cardiac hypertrophy. *Circ Res* 103: 1359-1362, 2008.

33. Monte E, Chen H, Kolmakova M, Parvatiyar M, Vondriska TM, Franklin S. Quantitative analysis of chromatin proteomes in disease. *J Vis Exp* 70: 4294-2012.

34. Monte E, Mouillesseaux K, Chen H, Kimball T, Ren S, Wang Y, Chen JN, Vondriska TM, Franklin S. Systems proteomics of cardiac chromatin identifies nucleolin as a regulator of growth and cellular plasticity in cardiomyocytes. *Am J Physiol Heart Circ Physiol* 305: H1624-H1638, 2013.

35. Nagandla H, Lopez S, Yu W, Rasmussen TL, Tucker HO, Schwartz RJ, Stewart MD. Defective myogenesis in the absence of the muscle-specific lysine methyltransferase SMYD1. *Dev Biol* 410: 86-97, 2016.

36. Olson EN, Backs J, McKinsey TA. Control of cardiac hypertrophy and heart failure by histone acetylation/deacetylation. *Novartis Found Symp* 274: 3-12, 2006.

37. Olson EN, Schneider MD. Sizing up the heart: Development redux in disease. *Genes Dev* 17: 1937-1956, 2003.

38. Paige SL, Thomas S, Stoick-Cooper CL, Wang H, Maves L, Sandstrom R, Pabon L, Reinecke H, Pratt G, Keller G, Moon RT, Stamatoyannopoulos J, Murry CE. A temporal chromatin signature in human embryonic stem cells identifies regulators of cardiac development. *Cell* 151: 221-232, 2012.

39. Pasumarthi KB, Field LJ. Cardiomyocyte cell cycle regulation. *Circ Res* 90: 1044-1054, 2002.

40. Rajabi M, Kassiotis C, Razeghi P, Taegtmeyer H. Return to the fetal gene program protects the stressed heart: a strong hypothesis. *Heart Fail Rev* 12: 331-343, 2007.

41. Sano M, Minamino T, Toko H, Miyauchi H, Orimo M, Qin Y, Akazawa H, Tateno K, Kayama Y, Harada M, Shimizu I, Asahara T, Hamada H, Tomita S, Molkentin JD, Zou Y, Komuro I. p53-induced inhibition of Hif-1 causes cardiac dysfunction during pressure overload. *Nature* 446: 444-448, 2007.

42. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* 13: 2498-2504, 2003.

43. Sohal DS, Nghiem M, Crackower MA, Witt SA, Kimball TR, Tymitz KM, Penninger JM, Molkentin JD. Temporally regulated and tissue-specific gene manipulations in the adult and embryonic heart using a tamoxifen-inducible Cre protein. *Circ Res* 89: 20-25, 2001.

44.   Srivastava D, Olson EN. A genetic blueprint for cardiac development. *Nature* 407: 221-226, 2000.

45.   Stewart MD, Lopez S, Nagandla H, Soibam B, Benham A, Nguyen J, Valenzuela N, Wu HJ, Burns AR, Rasmussen TL, Tucker HO, Schwartz RJ. Mouse myofibers lacking the SMYD1 methyltransferase are susceptible to atrophy, internalization of nuclei and myofibrillar disarray. *Dis Model Mech* 9: 347-359, 2016.

46.   Tan X, Rotllant J, Li H, De Deyne P, Du SJ. SmyD1, a histone methyltransferase, is required for myofibril organization and muscle contraction in zebrafish embryos. *Proc Natl Acad Sci USA* 103: 2713-2718, 2006.

47.   van Berlo JH, Maillet M, Molkentin JD. Signaling effectors underlying pathologic growth and remodeling of the heart. *J Clin Invest* 123: 37-45, 2013.

48.   Voelkel T, Andresen C, Unger A, Just S, Rottbauer W, Linke WA. Lysine methyltransferase Smyd2 regulates Hsp90-mediated protection of the sarcomeric titin springs and cardiac function. *Biochim Biophys Acta* 1833: 812-822, 2013.

49.   Wamstad JA, Alexander JM, Truty RM, Shrikumar A, Li F, Eilertson KE, Ding H, Wylie JN, Pico AR, Capra JA, Erwin G, Kattman SJ, Keller GM, Srivastava D, Levine SS, Pollard KS, Holloway AK, Boyer LA, Bruneau BG. Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage. *Cell* 151: 206-220, 2012.

50.   Wei JQ, Shehadeh LA, Mitrani JM, Pessanha M, Slepak TI, Webster KA, Bishopric NH. Quantitative control of adaptive cardiac hypertrophy by acetyltransferase p300. *Circulation* 118: 934-946, 2008.

51.   Wohlschlaeger J, Schmitz KJ, Schmid C, Schmid KW, Keul P, Takeda A, Weis S, Levkau B, Baba HA. Reverse remodeling following insertion of left ventricular assist devices (LVAD): A review of the morphological and molecular changes. *Cardiovasc Res* 68: 376-386, 2005.

52.   Zhang CL, McKinsey TA, Chang S, Antos CL, Hill JA, Olson EN. Class II histone deacetylases act as signal-responsive repressors of cardiac hypertrophy. *Cell* 110: 479-488, 2002.

53.   Zhang QJ, Chen HZ, Wang L, Liu DP, Hill JA, Liu ZP. The histone trimethyl lysine demethylase JMJD2A promotes cardiac hypertrophy in response to hypertrophic stimuli in mice. *J Clin Invest* 121: 2447-2456, 2011.

Chapter 2: The Histone Chaperone NAP1L4 Regulates Cardiac Transcription in Hypertrophy

Introduction

Histone chaperones are a key group of regulators involved in histone dynamics including nucleosome assembly and disassembly[1]. They are involved in various transcription-associated exchanges such as initiating robust transcription and elongation[2].

NAP1L4, also called NAP2, is a histone chaperone and was first identified from a human placenta cDNA library. This gene is ubiquitously and biallelically expressed homology of yeast nucleosome assembly protein (NAP1)[3]. Additional roles of NAP1L4 may be initiated by a stretch of amino acid residues in NAP1L4 at position 255-274 that has characteristics similar to those of nuclear localization signals (NLS)[4]. Recombinant NAP1L4 has the ability to bind linker and core histones alike[4]. For example, a study performing mass spectrometry analysis in vitro models of permeablized HeLa cell systems found NAP1L4 binding to histones H2A-H2B[5]. NAP1L4 is regulated by phosphorylation and found to be dephosphorylated following DNA damage stimuli[6] or during cell division. Cell cycle changes perturb phosphorylation status in NAP1L4; between G0 and G1 phases, phosphorylated NAP1L4 stays in the cytoplasm with core histones, whereas during the G1/S phase transition, NAP1L4 could be dephosphorylated and NLS was revealed, leading to nucleus transportation of NAP1L4 followed by histone deposition onto DNA[4, 7]. In addition to the aforementioned

nucleosome assembly mechanism, NAP1L4 can also act as a nucleosome disassembly protein.

NAP1L4 control via phosphorylation may not be circumscribed to cellular division, as it can be rapidly modified by mitogen-activated protein kinase signaling[8] that is involved in heart function and diseases[9]. A previous co-immunoprecipitation study has demonstrated that the association of p300-CBP coactivator family with NAP1L4, increasing the expression of p300 targets, including p53 and E2F[10]. Moreover, NAP1L4 was suggested to interact with topoisomerase I, a transcriptional critical enzyme that loosens nucleosomes for increased accessibility[11, 12]. In terms of spatial regulation, NAP1L4 can bind to diacylglycerol kinase ζ (DGKζ)--the dominating DGK isoform in ventricular myocardium[13]--preventing the translocation of DGKζ into the nucleus[14]. The relocation of DGKζ has been shown to be important during ischemia/reperfusion[15].

Our previous proteomic experiments identified altered expression of NAP1L4 in the hypertrophic and failing mouse heart after pressure overload, raising the possibility that this protein was an unrecognized participant in disease-associated chromatin remodeling. In the present study, we used PHE to induce hypertrophy in NAP1L4 knockdown NRVMs to investigate the mechanisms by NAP1L4 inhibits the activity of the fetal gene program (a group of genes previously shown to be regulated in the setting of of hypertrophy) in cardiomyocytes. Overexpression of NAP1L4 was also examined to further explore the protein's function in cardiomyocytes.

Material and Methods

NAP1L4 knockdown with PHE treatment using NRVM

NRVMs were acquired from the lab of Dr. Yibin Wang by enzymatic dissociation from the ventricles of 1-day-old pups and plated in serum media for the first 24h, after which the NRVMs were cultured in serum- and antibiotic-free media containing 1% insulin-transferrin-sodium selenite supplement (time 0). Four groups were designed: 1) the control group with Lipofectamine RNAi Max (Lipo) only (Invitrogen, 13778-150); 2) siRNA knockdown group transfected with 50nM NAP1L4 siRNA (Qiagen, SI01797901) and corresponding Lipo; 3) PHE only group (10µM) treated at time 48h; 4) siRNA knockdown with PHE treatment at time 48h (Figure 2-1). After a treatment totaling 96h, NRVMs were harvested and followed by western blot, qPCR, and cell size measurement through Nikon eclipse TE2000-U microscope and quantified by Image J. Western blot was implemented by an anti-NAP1L4 antibody (ab21631, Abcam). Primers [forward (F) and reverse (R)] that used in this study were as follows: NAP1L4 F-5'-GCCGCCGAACTGTGTGGGAA-3' R-5'-TGCAACAACCCACCCACCCC-3'; GAPDH F-5'- CCCACTAACATCAAATGGGG-3' R-5'-CCTTCCACAATGCCAAAGTT-3'; ANF F-5'-CTGATGGATTTCAAGAACCTGCT-3' R-5'-CTCTGGGCTCCAATCCTGTC-3'; SERCA2a F-5'-CCTTCTACCAGCTGAGTCATTT-3' R-5'-CAGATGGAGCCCACGACCCA-3'; α-MHC F-5'-GAACAGCTGGGAGAAGGGGG-3' R-5'-GCCTCTGAGGCTATTCTATTGG-3'; and β-MHC F-5'-CTCAACTGGGAAGAGCATCCA-3' R-5'-CCTTCAGCAAACTCTGGAGGC-3'.

NAP1L4 overexpression with adenovirus

An adenovirus for NAP1L4 with GFP was ordered from Vector BioLabs and transfection of NRVMs performed as the experiment group. The empty GFP (Ad-CMV-

GFP) was used as a control (Vector BioLabs, 1060). After the 96h of NAP1L4 overexpression, cells were harvested and western blot, qPCR, and cell size calculations were completed as previously described.

Results

NAP1L4 knockdown prevents PHE-induced hypertrophy

After knocking down NAP1L4 in normal NRVMs and PHE-treated NRVMs, the size of NRVMs decreased significantly in siRNA knockdown NRVMs compared with control (Lipo). Meanwhile, the cell size of PHE-treated NAP1L4 knockdowns decreased when compared to PHE-treated NRVMs (scrambled RNA was used as control and was similar to Lipo, data not shown) (Figure 2-2A). Meanwhile, with respect to the control (Lipo) group, cells in the PHE-treated group showed a marked increase in cell size, highlighting the effectiveness of agonist. However, following NAP1L4 knockdown, the cell size was not significantly increased compared to the control group (Figure 2-2B).

The western blot of the NAP1L4 for different groups revealed that NAP1L4 inhibition via siRNA was effective—we observed decreased expression level of NAP1L4 in the siRNA knockdown group. Also, in the PHE-treated group, there was an increase of NAP1L4, indicating the potential function of NAP1L4 in the diseased condition (Figure 2-2C).

NAP1L4 knockdown inhibits fetal gene reprogramming

The activity of the fetal gene program constituents were measured for all experimental groups. Comparing changes in gene expression between the PHE-treated

and control groups, trial natriuretic factor (ANF) transcript levels increased significantly. Conversely, sarco/endoplasmic reticulum Ca2+-ATPase (SERCA) and α-myosin heavy chain (αMHC) transcription drastically decreased. The results show that the fetal gene program was activated during the PHE treatment. However, when first knocking down NAP1L4 from cells and then subsequently treating them with PHE, the expression of ANF, SERCA and αMHC exhibited a reversed trend compared to the just PHE-treated group. This validates the inhibition of the fetal gene reprogramming induced by NAP1L4 knockdown (Figure 2-3).

NAP1L4 overexpression increases cell size

We first successfully overexpressed NAP1L4 in NRVMs (Figure 2-4A). At the same time, we compared the cell size of the control group, the GFP-control group, and NAP1L4 overexpression groups (Figure 2-4B-C). We found that the cell size of NAP1L4 overexpressed groups was significantly increased compared to both control and GFP-control groups, indicating the cell size increase may have been due to the overexpression of NAP1L4.

Discussion

NAP1L4 is a histone chaperone that has been identified by our lab to be up-regulated in hypertrophic and failing mouse hearts compared to control mice using mass spectrometry (MS)[16]. After a 48h PHE treatment revealed an increase of NAP1L4 consistent with the MS results, as confirmed by western blot. Furthermore, knockdown of NAP1L4 reversed the cell size increase induced by PHE treatment. At the same time,

NAP1L4 knockdown inhibited the fetal gene reprogramming—shown in hypertrophic or failing hearts—that could be induced by PHE treatment. Also, overexpression of NAP1L4 depicted an increase in the cell size of NRVMs relative to the GFP control. The culmination of these results supports the idea that NAP1L4 may have a cardioprotective role and could be a potential clinical target for the treatment of hypertrophy.

In this study, we used NRVMs in lieu of adult cardiomyocytes because they are genetically manipulated readily and, to a certain degree, exhibit the hypertrophic signaling seen in adult cells. Adult cardiomyocytes, on the other hand, are difficult to transfect and are not conductive to time course studies due to dedifferentiation. Finally, the whole heart for this experiment could be confounded by the integration of non-cardiomyocytes cells.

The future directions for this study are to first understand the potential signal pathways and targets of NAP1L4. These could be achieved by sequencing the RNA for differentially expressed genes comparing the control group, PHE-treated group, NAP1L4 knockdown groups, both with and without PHE treatment. Chromatin immunoprecipitation sequencing (ChIP) or ChIP-qPCR could be utilized to find the possible targets of NAP1L4. Finally, applying our experimental workflow to adult cardiomyocytes and mouse model could confirm the function of NAP1L4, *in vivo*.
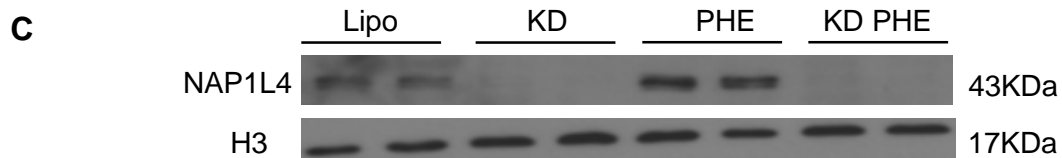
Figure 2-1. NAP1L4 knockdown workflow. Four experimental groups were designed: Lipo/control group, NAP1L4 siRNA KD group, PHE-treated group, and NAP1L4 KD with PHE treatment. Control cells were treated with Lipo only. The agonist phenylephrine (10μM) was added to NRVMs at 48h in the timeline and cells were harvested at 96h.
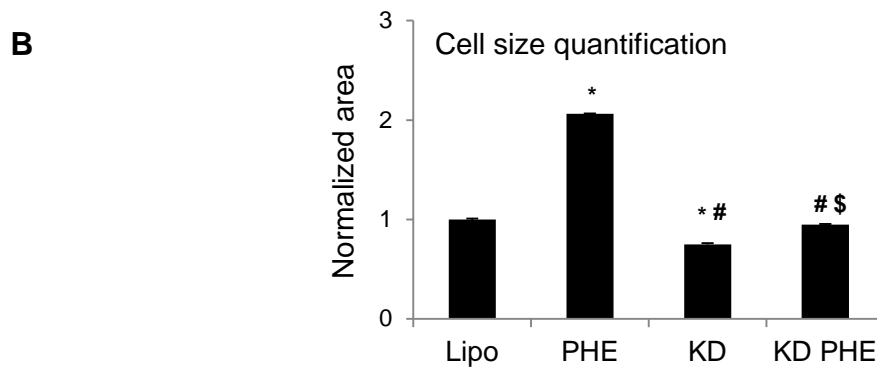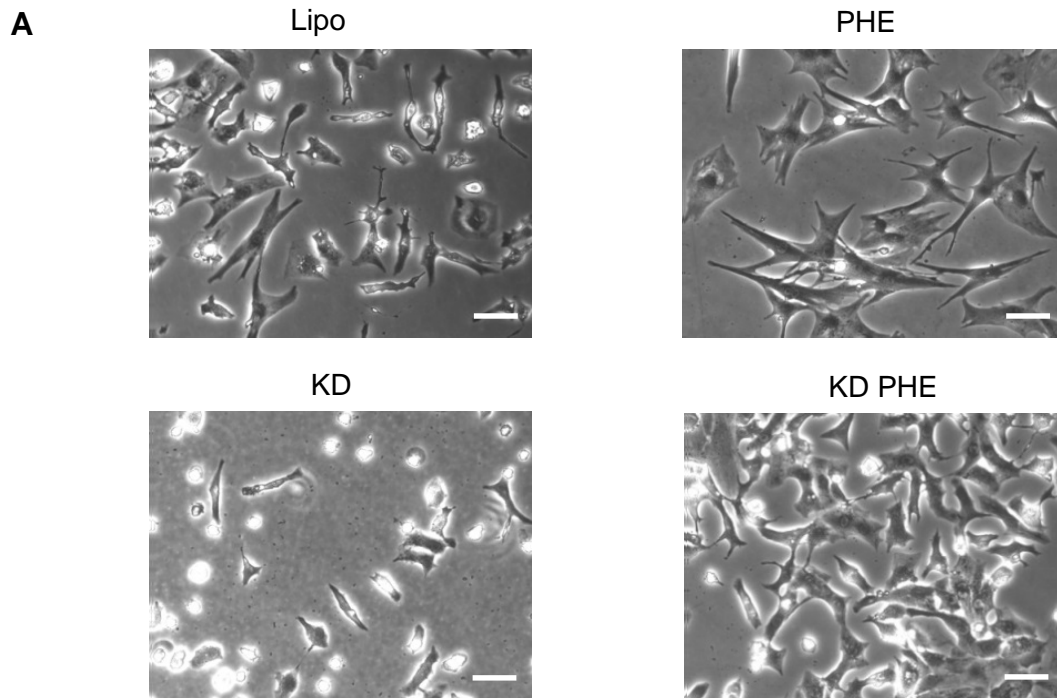
Figure 2-2. NAP1L4 knockdown prevents PHE-induced cell size increase. A. Bright-field images show representative NRVMs used for size quantification. **B.** Cell sizes were quantified using 100 cells of each condition group (2 replicates). **C.** Western blot (2 replicates) for NAP1L4 and H3 as the loading control was performed. Scale bar: 50μm. Scrambled RNA was used as an additional control and was similar to Lipo, data not shown. * P<0.05 vs Lipo; # P<0.05 vs PHE; $ P<0.05 vs KD
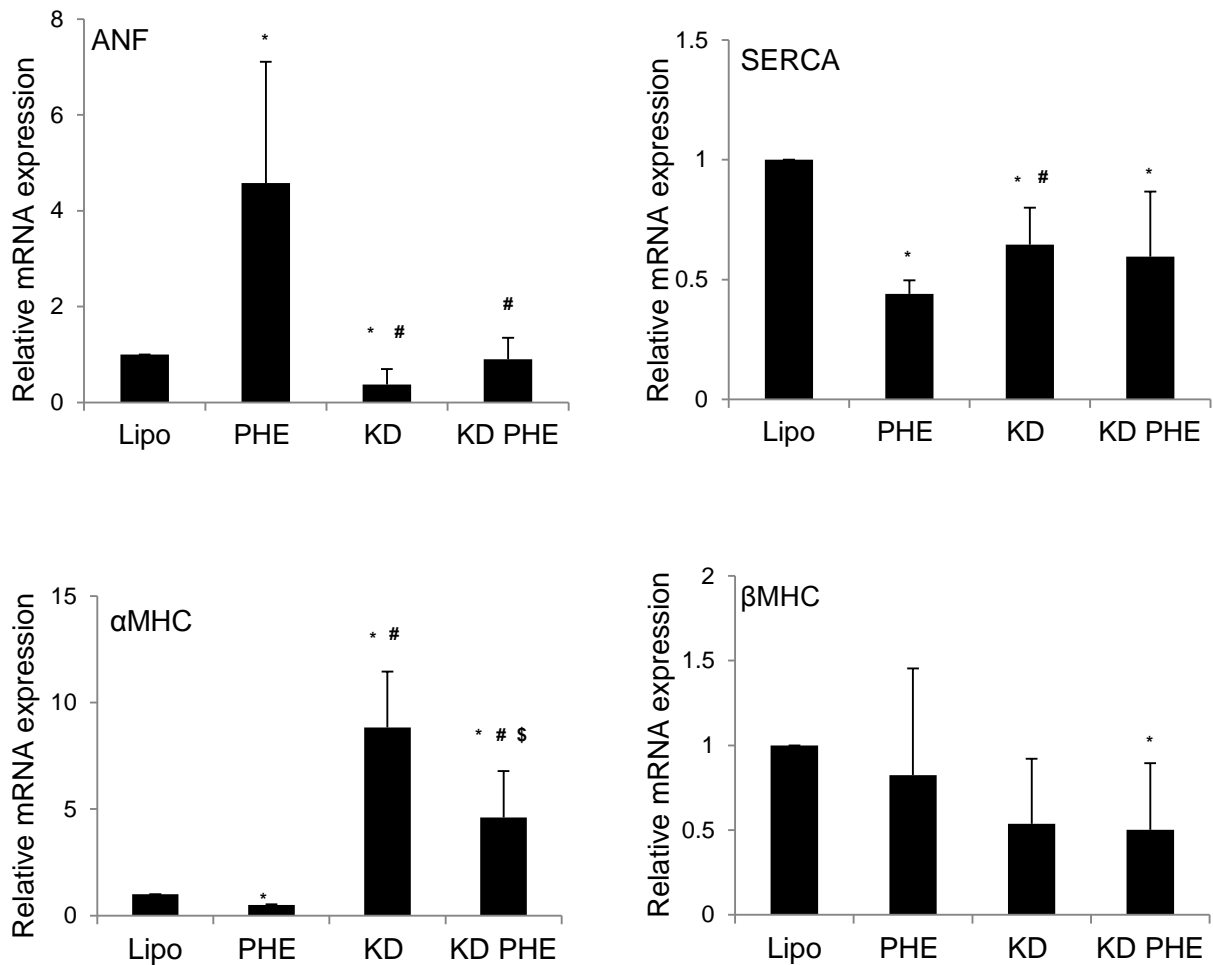
Figure 2-3. NAP1L4 knockdown prevents PHE-induced fetal gene reactivation.
Fetal genes ANF, SERCA, αMHC and βMHC were tested after NAP1L4 KD for 96h
using qPCR (4 replicates). Scrambled RNA was used as a control and was similar to
Lipo, data not shown. * P<0.05 vs Lipo; # P<0.05 vs PHE; $ P<0.05 vs KD.

A

B

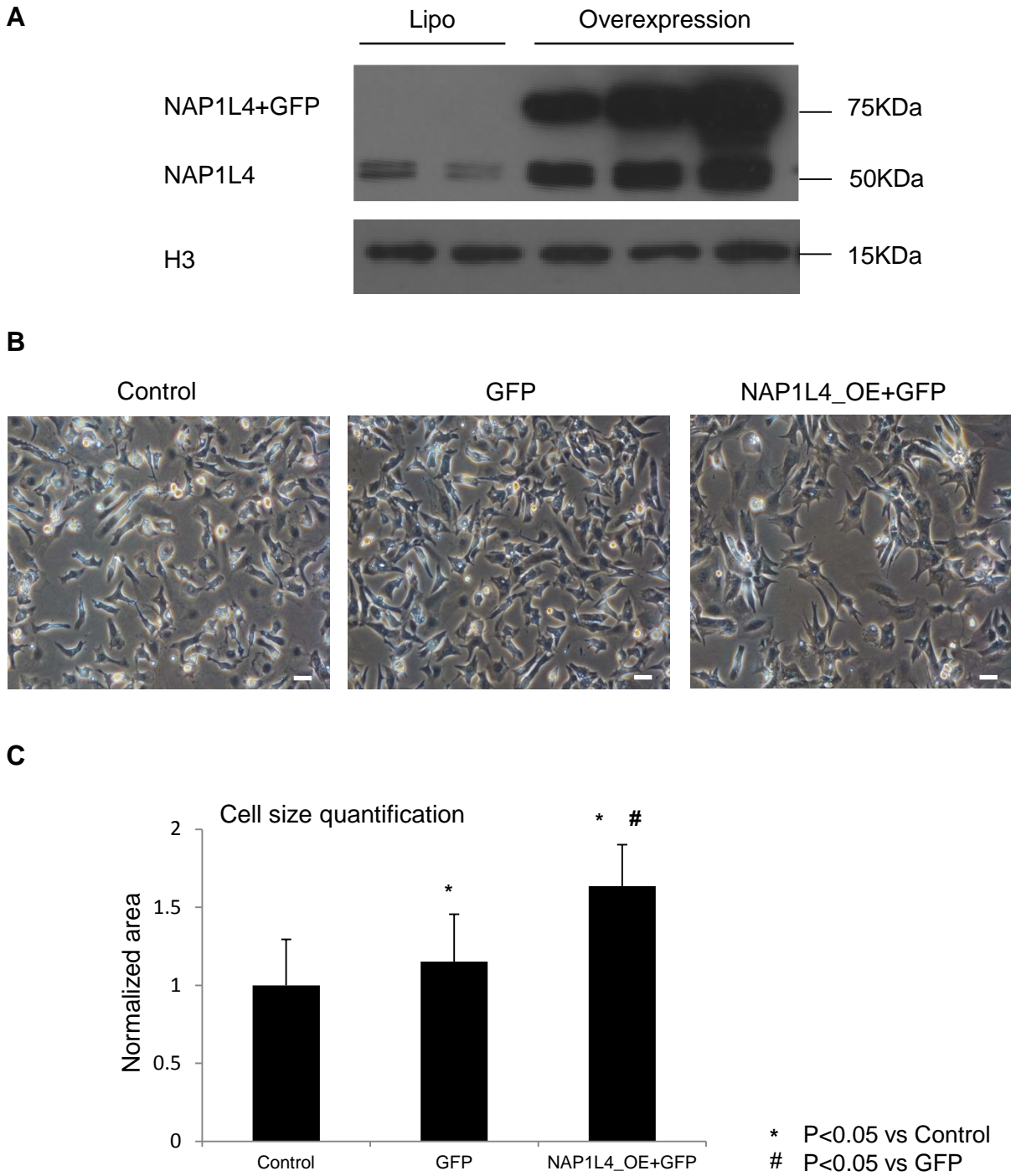| Control | GFP | NAP1L4_OE+GFP |

C

* P<0.05 vs Control
# P<0.05 vs GFP

Figure 2-4. NAP1L4 overexpression increases the cell size. A. The western blot of NAP1L4 overexpression. **B.** Bright-field images show representative NRVMs used for size quantification. **C.** Cell sizes were quantified for each condition. Scale bar: 50μm. * P<0.05 vs Control; # P<0.05 vs GFP

References

1. D'Arcy S, Martin KW, Panchenko T, Chen X, Bergeron S, Stargell LA, Black BE, Luger K. Chaperone nap1 shields histone surfaces used in a nucleosome and can put h2a-h2b in an unconventional tetrameric form. *Molecular cell*. 2013;51:662-677

2. Venkatesh S, Workman JL. Histone exchange, chromatin structure and the regulation of transcription. *Nature reviews. Molecular cell biology*. 2015;16:178-189

3. Hu RJ, Lee MP, Johnson LA, Feinberg AP. A novel human homologue of yeast nucleosome assembly protein, 65 kb centromeric to the p57kip2 gene, is biallelically expressed in fetal and adult tissues. *Human molecular genetics*. 1996;5:1743-1748

4. Rodriguez P, Munroe D, Prawitt D, Chu LL, Bric E, Kim J, Reid LH, Davies C, Nakagama H, Loebbert R, Winterpacht A, Petruzzi MJ, Higgins MJ, Nowak N, Evans G, Shows T, Weissman BE, Zabel B, Housman DE, Pelletier J. Functional characterization of human nucleosome assembly protein-2 (nap1l4) suggests a role as a histone chaperone. *Genomics*. 1997;44:253-265

5. Kimura H, Takizawa N, Allemand E, Hori T, Iborra FJ, Nozaki N, Muraki M, Hagiwara M, Krainer AR, Fukagawa T, Okawa K. A novel histone exchange factor, protein phosphatase 2cgamma, mediates the exchange and dephosphorylation of h2a-h2b. *The Journal of cell biology*. 2006;175:389-400

6. Dirksen EH, Pinkse MW, Rijkers DT, Cloos J, Liskamp RM, Slijper M, Heck AJ. Investigating the dynamic nature of the interactions between nuclear proteins and histones upon DNA damage using an immobilized peptide chemical proteomics approach. *Journal of proteome research*. 2006;5:2380-2388

7. Rodriguez P, Pelletier J, Price GB, Zannis-Hadjopoulos M. Nap-2: Histone chaperone function and phosphorylation state through the cell cycle. *Journal of molecular biology*. 2000;298:225-238

8. Lewis TS, Hunt JB, Aveline LD, Jonscher KR, Louie DF, Yeh JM, Nahreini TS, Resing KA, Ahn NG. Identification of novel map kinase pathway signaling targets by functional proteomics and mass spectrometry. *Molecular cell*. 2000;6:1343-1354

9. Rose BA, Force T, Wang Y. Mitogen-activated protein kinase signaling in the heart: Angels versus demons in a heart-breaking tale. *Physiological reviews*. 2010;90:1507-1546

10. Shikama N, Chan HM, Krstic-Demonacos M, Smith L, Lee CW, Cairns W, La Thangue NB. Functional interaction between nucleosome assembly proteins and p300/creb-binding protein family coactivators. *Molecular and cellular biology*. 2000;20:8933-8943

11. Rodriguez P, Ruiz MT, Price GB, Zannis-Hadjopoulos M. Nap-2 is part of multi-protein complexes in hela cells. *Journal of cellular biochemistry*. 2004;93:398-408

12. Collins I, Weber A, Levens D. Transcriptional consequences of topoisomerase inhibition. *Molecular and cellular biology*. 2001;21:8437-8451

13. Takeda M, Kagaya Y, Takahashi J, Sugie T, Ohta J, Watanabe J, Shirato K, Kondo H, Goto K. Gene expression and in situ localization of diacylglycerol kinase isozymes in normal and infarcted rat hearts: Effects of captopril treatment. *Circulation research*. 2001;89:265-272

14. Okada M, Hozumi Y, Ichimura T, Tanaka T, Hasegawa H, Yamamoto M, Takahashi N, Iseki K, Yagisawa H, Shinkawa T, Isobe T, Goto K. Interaction of nucleosome assembly proteins abolishes nuclear localization of dgkzeta by attenuating its association with importins. *Experimental cell research*. 2011;317:2853-2863

15. Akiyama H, Hozumi Y, Nakano T, Kubota I, Goto K. Nuclear relocation of dgkzeta in cardiomyocytes under conditions of ischemia/reperfusion. *Histology and histopathology*. 2011;26:1383-1390

16. Monte E, Mouillesseaux K, Chen H, Kimball T, Ren S, Wang Y, Chen JN, Vondriska TM, Franklin S. Systems proteomics of cardiac chromatin identifies nucleolin as a regulator of growth and cellular plasticity in cardiomyocytes. *American journal of physiology. Heart and circulatory physiology*. 2013;305:H1624-1638

Chapter 3: The Role of Circular RNA in Cardiac

Hypertrophy and Failure

Introduction

Non-coding RNAs are RNAs that do not translate into proteins. CircRNAs are an emerging type of non-coding RNAs that exhibit abundant, albeit tissue-specific expression. RNA sequencing data analysis of five different mouse tissues--including heart--showed that 20% of protein-coding genes produced circRNAs[1]. Unlike linear RNAs, circRNAs are formed through a process known as back-splicing, featuring a covalently joined loop structure which can be detected by divergent primers that face away from each other. The formation and function of circRNAs are still poorly understood. It has been shown that the RNA-binding protein QKI (protein quaking) could promote circRNA biogenesis during human epithelial-mesenchymal transition[2]. circRNAs can function as a microRNA sponge[3], regulating transcription of corresponding linear forms of circRNAs[4], and serving as a marker for diseases[5]. In the heart, only a few studies have revealed that circRNAs decoy microRNA and proteins, alter mRNA expression, or serve as biomarkers. The first circRNA detected in cardiovascular diseases is located in the antisense of INK locus (ANRIL) that correlated with INK/ARF expression and atherosclerotic risk[6]. The mechanisms by which circRNAs impact heart disease are not fully understood.

In this study, we isolated cardiomyocytes from mice that underwent transverse aortic constriction (TAC), followed by Ribo-Zero RNA sequencing to obtain a landscape of circRNAs in the cardiomyocyte. Additionally, we sought to confirm the existence of

cardiac-related circRNAs using quantitative polymerase chain reaction (qPCR) and electrophoresis followed by Sanger sequencing. Finally, we continued our exploration into circRNAs by understanding the potential regulatory function of circMyocd through its linear counterpart. siRNA knockdown of circMyocd was performed and linear Myocd expression level was quantified.

Materials and Methods

Transverse aortic constriction and cardiomyocyte isolation

All animals used in this study were approved by the University of California, Los Angeles (UCLA) and supported by the university Animal Research Committee. Six adult C57BL/6 male and female mice aged 8 weeks were randomly, but evenly dispersed into one of either the sham group or the TAC group. All mice underwent a similar procedure, but aortic clamps were only placed on the TAC group mice. The detailed procedures for TAC were previous described[7] and carried out by our long-term collaborator Dr. Yibin Wang. Heart failure was confirmed by a cardiac function examination via echocardiogram and calculated heart-to-body weight ratios. After 4 weeks of surgery, all animals were treated with heparin to prevent blood coagulation, and subsequently sacrificed. Left ventricular cardiomyocytes were isolated using the Langendorff method[8]. Tyrode's solution with collagenase type-II and protease type-XIV was applied to the myocytes. Cells were successfully isolated at the level wherein overall cellular composition was over 95% cardiomyocytes[9].

Ribo-Zero RNA sequencing

Whole RNA was extracted from cardiomyocytes isolated from left ventricles by Qiagen RNeasy Mini Kit (Qiagen, 74104), following manufacturer instruction. Ribo-Zero RNA (Ribosomal RNA depletion) was obtained by Illumina Ribo-Zero rRNA Removal Kit (Illumina, MRZH11124). RNA-seq libraries were built by the UCLA Microarray Core facility and subsequently sequenced on an Illumina Hiseq 2000 by the UCLA Genetics Core facility. The condition for the RNA sequencing was set at 100 base pair, paired-end and non-strand specific, as previously described[9].

Bioinformatics for mRNA sequencing and circRNAs detection

After libraries were demultiplexed, fastq files for all 6 samples were obtained. For mRNA sequencing analysis, all reads were mapped to the mouse reference genome (mm10) using STAR[10]. SAMtools[11] was subsequently applied for sorting and htseq-count[12] was used for counting the number of mRNA reads. The DESeq2[13] package was used for differentially expressed gene calculations, significance here set by p-values below 0.05. A gene set enrichment analysis (GSEA)[14] was employed for gene pathways detection that based on ranking information of the differentially expressed genes not p-value.

For circular RNA detection, BWA-MEM[15] was utilized to map each fastq file to the reference genome (mm10). CIRI v2.0.1[16] was then used for de novo circRNA detection by Perl. The coordinate for each circRNA found by CIRI was elucidated using a vertical bar '|' that indicates the back-spliced junction site for each circRNA (Figure 3-1).

Primer design for mouse and rat genomes

For mouse genomes, primers for circular RNA were designed specifically to cover the back-spliced junctions. One exon on each side of the back-spliced junction was used for circRNA primer design (Figure 3-2). For verification, 3 forward primers (F1, F2, F3) and 3 reverse primers (R1, R2, R3) were designed, resulting in a total of 9 primers for each circRNA (F1R1, F1R2, F1R3, F2R1, F2R2, F2R3, F3R1, F3R2, F3R3). Meanwhile, 2 linear primers were designed to examine the existence of the linear RNA (Figure 3-2, Table 3-1).

The same method for generation of primers was used for circRNAs in rat genomes (Table 3-2). For the following qPCR experiments using NRVMs for circMyocd, we selected the F1 and R2 primers (Table 3-2). For qPCR experiments using NRVMs for linear Myocd, two sets of primers were applied. One primer set was designed based on the DNA sequencing that covered the linear Myocd but not the circMyocd. The other set of primer was revised from a previously published study[17]. The two linear Myocd primer sequences were as follow: linear Myocd_1: F: GGCTTTCCCCAGATCAGGCT; R: CCAGAGGAGCCTCAGTGGATT; linear Myocd_2: F: CAACTGTCACCTTTCCTGTCACG; R: CTACCAGCATCTTGTCCTTCTC.

RNA extraction, qPCR, electrophoresis, and Sanger sequencing

For mouse circRNA confirmation, RNA was extracted from mouse hearts using TRIzol (Thermo Fisher Scientific, 15596018), following manufacturer protocol. This was converted to cDNA by iScript cDNA Synthesis Kit (Bio-Rad, 170-8891). For each primer set, qPCR was performed using SsoFast EvaGreen Supermix (Bio-Rad, 172-5201).

Electrophoresis was done using 1% agarose gel for 30min at 120V for qPCR products. If there was only one clear band found at the targeted molecular weight of circRNAs, we selected one out of nine primers and submitted it Sanger sequencing at Laragen. We confirmed the existence of each circRNA if by ensuring the Laragen sequencing covered the back-spliced junction and that the signal for each nucleotide sequenced was distinct.

circRNA knockdown

Two circMyocd siRNAs against rat genomes were designed to target the back-spliced junction from GE Healthcare Dharmacon. dTdT overhangs were added to circMyocd siRNA in order to enhance nuclease resistance under experimental condition[18]. RNase-free water was added to siRNAs to make the final concentration 20µM. The two circMyocd sequences were as follow:

1) Sense: 5' GGCUAUAAAAGUCUUACAGUUdTdT 3';

   Antisense: 5' AACUGUAAGACUUUUAUAGCCdTdT 3'.

2) Sense: 5' AGGCUAUAAAAGUCUUACAGUdTdT 3';

   Antisense: 5' ACUGUAAGACUUUUAUAGCCUdTdT 3'.

NRVMs were utilized for the knockdown experiment, which was performed using 12.5µl of circMyocd siRNA in 0.5ml of Opti-MEM (Gibco, 31985070) combined with 15ul Lipofectamine™ RNAiMAX Transfection Reagent (Thermo fisher scientific, 13778075) in 0.5ml of Opti-MEM. The control experiment was performed similarly but without circMyocd siRNA. The siRNA Duplex was applied as a negative control (Qiagen, 1027310).

Results

mRNA analysis confirmed heart hypertrophy after TAC construction

Principle component analyses of RNA-seq data clearly separated transcriptomes from TAC and control groups (data not shown). After analyzing the differential gene expression changes by DESeq2, there were 1688 genes that were significantly down-regulated and 1728 genes that were significantly up-regulated in TAC conditions. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways indicated many cardiac related biological process and pathways. The GSEA analysis, which represented all differentially expressed genes, confirmed the presence of expression levels changing in genes related to cardiac diseases as a result of TAC stress (Figure 3-3).

Cardiac-related circRNAs confirmation in mice and rat genome

Based on the output of CIRI, a total of 941 circular RNAs were detected. We selected and confirmed 13 circRNAs by qPCR, electrophoresis, and Sanger sequencing including: circMyocd (chr11:65218529|65233184), circAtp5c1 (chr2:10058996|10064311), circPpp2r3a (chr9:101125023|101153874), circZfp148 (chr16:33421321|33434974) and 3 circRyr2 (chr13:11759671|11785141, chr13:11680966|11688013 and chr13:11769852|11785141) and 6 circTtn (chr2:76857904|76881839, chr2:76857904|76884197, chr2:76863274|76868085, chr2:76802231|76818816, chr2:76847388|76881839, chr2:76847388|76884197) (Figure 3-4). We concluded that certain positions on the genome may have higher chances of forming circRNAs such as Ttn at chr2:76857904 and Ryr2 at chr13:

11785141. Of particular note is the experimental support of the concept of cardiac-related circular RNAs. Ttn (Titin) encodes a variety of proteins in the striated muscle. Ryr2 (Ryanodine Receptor 2) was found in cardiac muscle sarcoplasmic reticulum. Myocd (Myocardin) is expressed in the heart, aorta, and smooth muscle. For the rat genome, we confirmed 4 rat counterparts including 2 circRyr2, circZfp148 and circMyocd (Figure 3-5).

Expression of circRNA for Myocardin was opposite the linear RNA counterpart

To understand the potential function of circRNA, we selected circMyocd for the reasons that it was confirmed in both mouse and rat genomes, and that there exists a human homolog based on the genome sequencing information[19], indicating conservation of circMyocd across species. Meanwhile, there was a decreasing trend of circMyocd in the TAC group when comparing the expression levels of sham and TAC conditions, posing a potential therapeutic target.

We designed two siRNAs against circMyocd at the back-spliced junction site and confirmed the effectiveness of circMyocd knockdown in NRVMs by qPCR (Figure 3-6A-C). Finally, we analyzed the expression level of linear Myocd using two different primers and found the expression level of the linear counterparts exhibited an increased trend, inverse to that of circMyocd (Figure 3-6D-G).

Discussion

In this study, Ribo-Zero RNA sequencing was used to examine the circular RNA transcriptome in the mouse heart and to examine changes during disease.

Echocardiography and mRNA analysis verified the heart failure condition in TAC mice. For the circRNA analysis, 941 circRNAs were found and 13 circRNAs including some cardiac-related circRNA were confirmed including circMyocd, 3 circRyr2 and 6 circTtn. Based the back-spliced junction coordinates of confirmed circRNAs, certain positions may have a higher chance of forming circRNAs and can produce different circRNAs. To understand the function of circRNA, we utilized rat cell model NRVMs and siRNA knockdown experiments for circMyocd. We found the expression of linear Myocd increase while the expression level of circMyocd was decreased. The negative correlation of linear and circular RNAs was also shown in previously research[4]. Meanwhile, other studies found that the expression level of linear and circular counterparts were independent of one another[1, 20], suggesting that the characteristics of circRNAs are unique to certain degrees.

Future directions for this study would encompass defining the function of circMyocd, particularly overexpressed circMyocd and cross-examine with the expression levels of linear Myocd to better confirm the potential negative correlation between circular and linear forms. At the same time, it would be interesting to analyze the protein levels of Myocd and the gene expression targets of the Myocd protein transcription factor activity after knockdown of circMyocd. Thus, we would have a more holistic understanding of the function of circMyocd. The designing of mouse models could further increase our understanding of the possible functions of circMyocd *in vivo*.
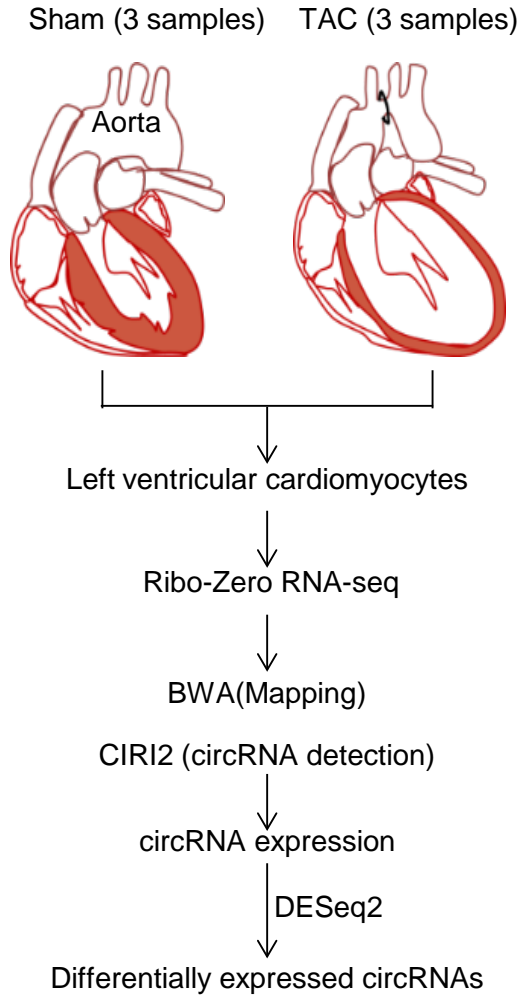
Sham (3 samples)   TAC (3 samples)

Aorta

Left ventricular cardiomyocytes

Ribo-Zero RNA-seq

BWA(Mapping)

CIRI2 (circRNA detection)

circRNA expression

DESeq2

Differentially expressed circRNAs

A

Linear primer

| Exon | Exon | Exon |

circular primer

B   Back-spliced junction

Junction read

Non-junction read

Figure 3-1. Ribo-Zero RNA-seq workflow. RNA-seq was performed for left ventricular cardiomyocytes from 3 sham and 3 TAC mice. CIRI2 was used for circRNA detection and DESeq2 was applied to find differentially expressed circRNAs.
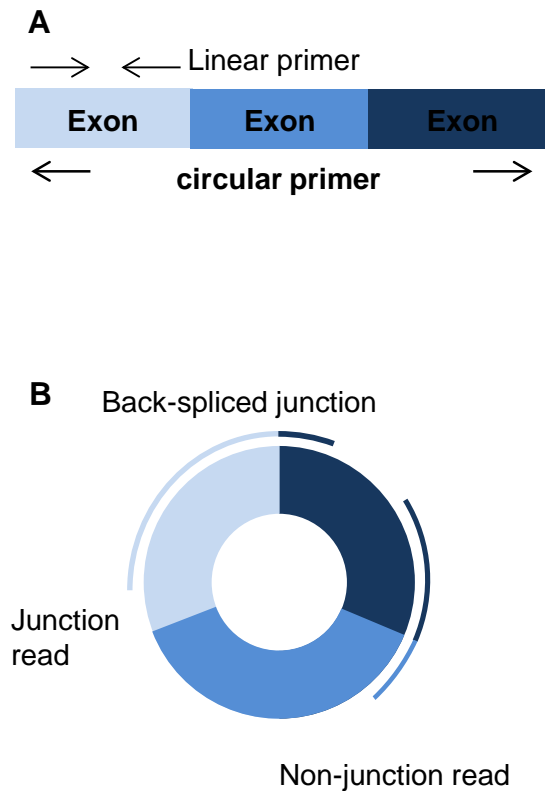
Figure 3-2. Schematic view of circRNA. A. linear mRNA. The divergent primer for circRNAs is shown. **B.** Junction read covers the back-spliced junction.
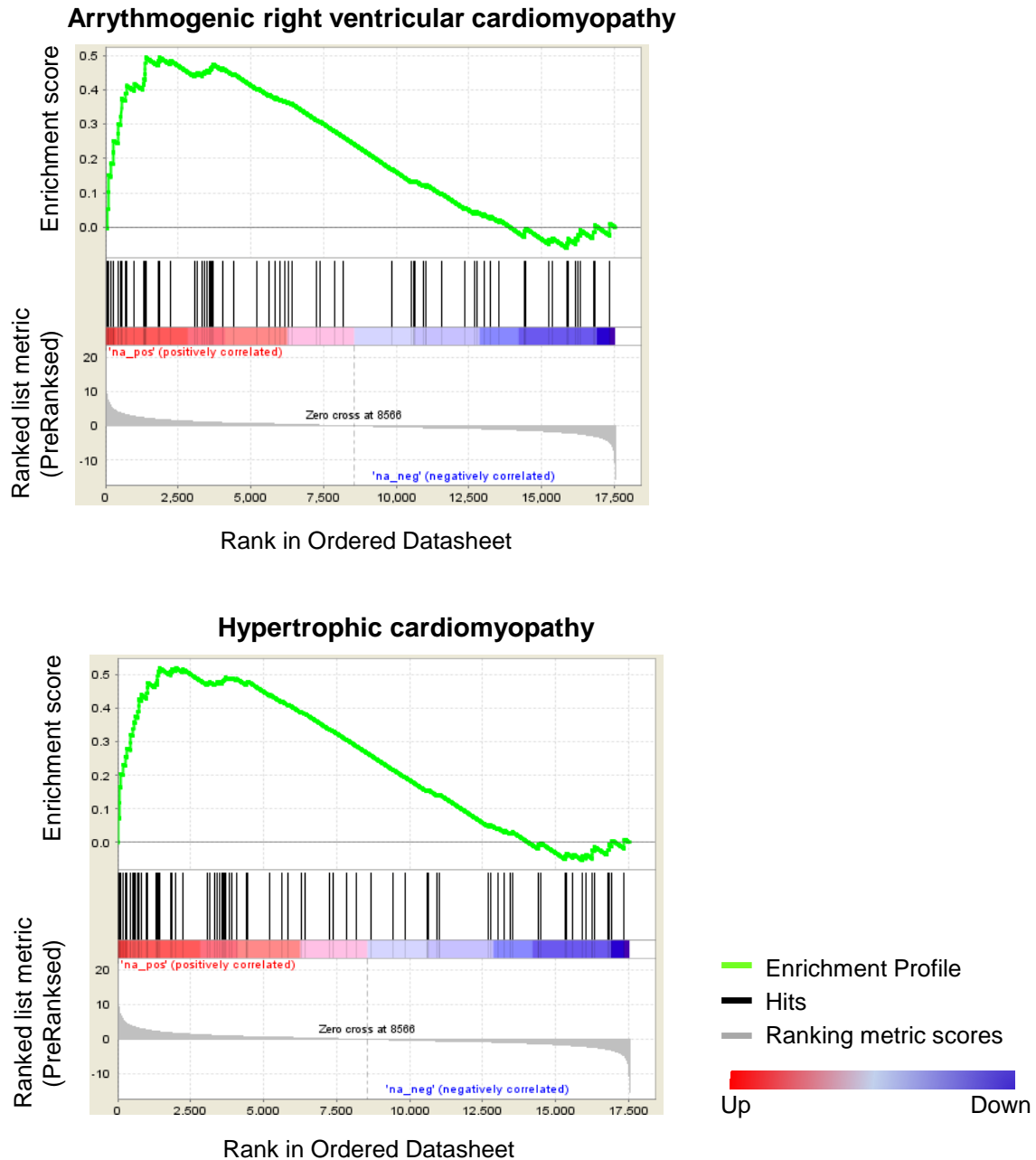
Arrythmogenic right ventricular cardiomyopathy

Rank in Ordered Datasheet

Hypertrophic cardiomyopathy

Rank in Ordered Datasheet

Enrichment Profile
Hits
Ranking metric scores

Up   Down

Figure 3-3. The gene expression profile generated by GSEA using the rank information after control and TAC mice comparison. These two terms were cardiac-related and among the top 50 GSEA results. The green line is the enrichment profile indicates the density of genes in the sequencing data (top part of the figure). Every line in the middle part represents a gene in the dataset. The red bar revealed the gene were up-regulated under TAC condition and the blue bar conveyed the genes in this region were down-regulated under TAC conditions.
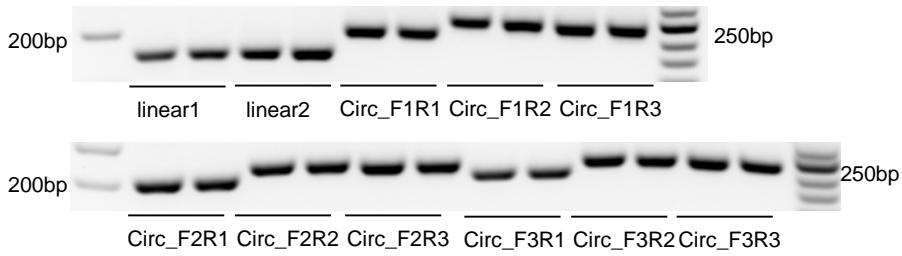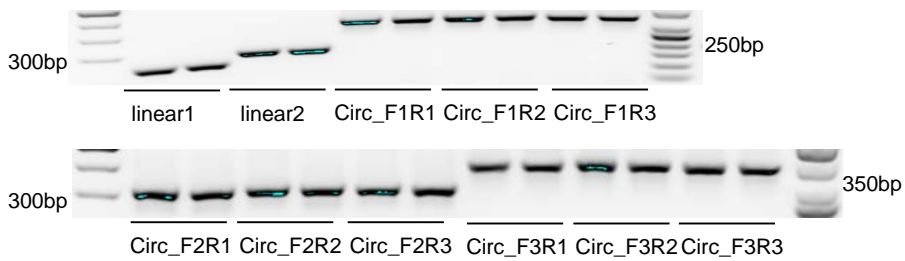
circTtn_chr2:76857904|76881839
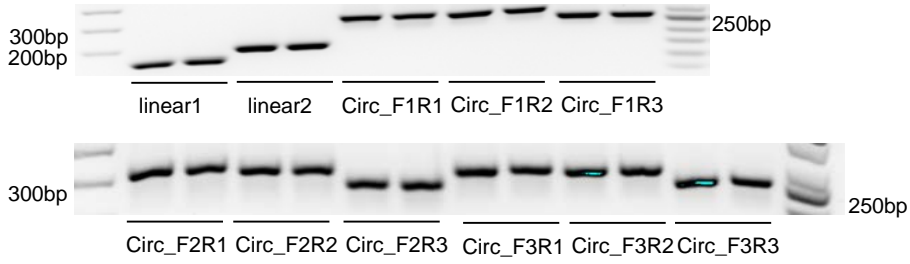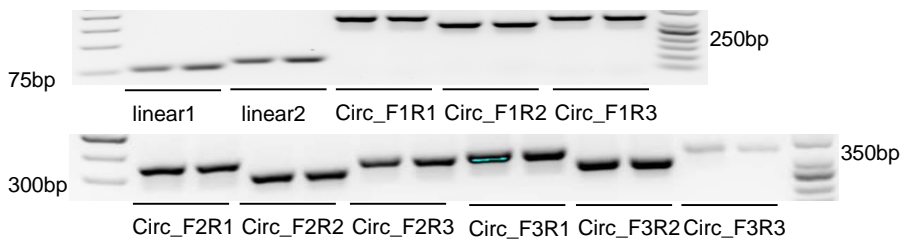

circRyr2_chr13:11759671|11785141
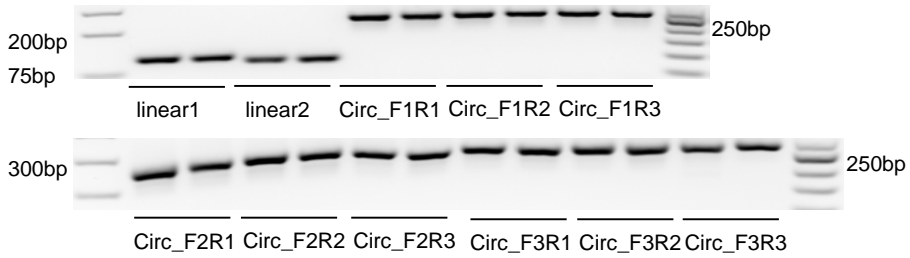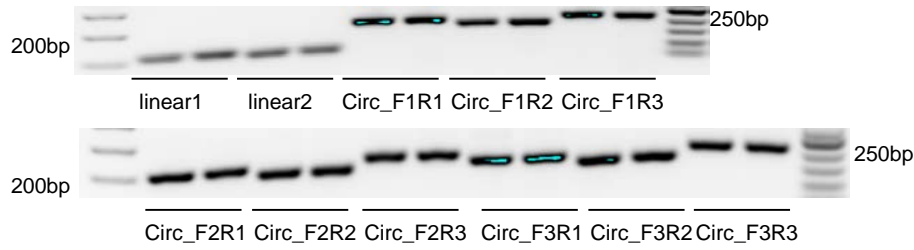

circMyocd_chr11:65218529|65233184

circTtn_chr2:76863274|76868085

circTtn_chr2:76857904|76884197

circRyr2_chr13:11680966|11688013

circRyr2_chr13:11769852|11785141
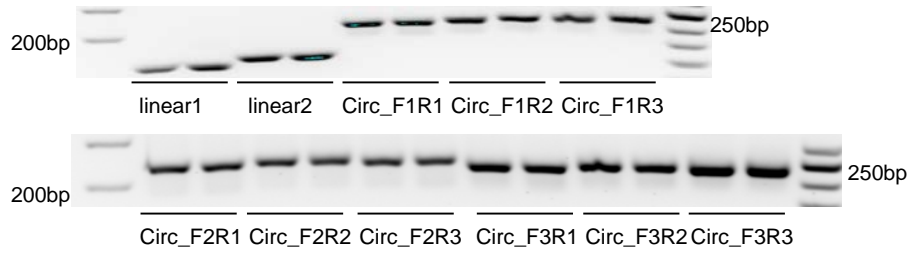
circAtp5c1_chr2:10058996|10064311
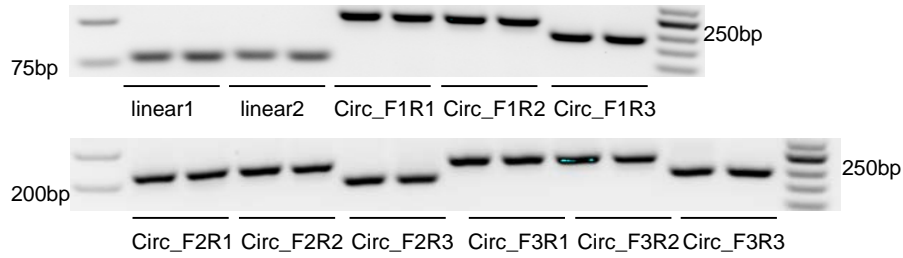
circPpp2r3a_chr9:101125023|101153874
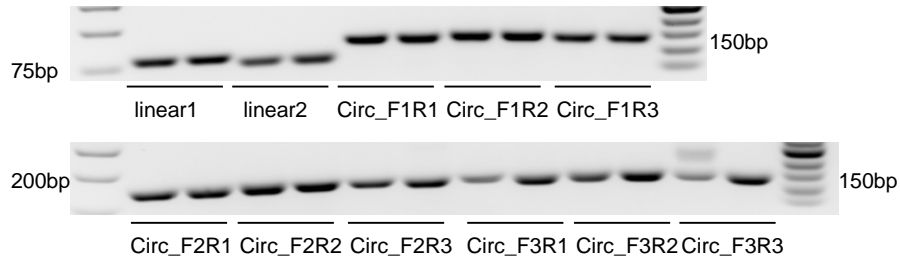
circZfp148_chr16:33421321|33434974
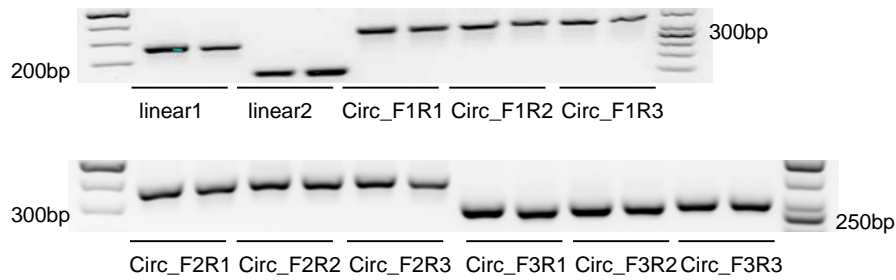
circTtn_chr2:76802231|76818816
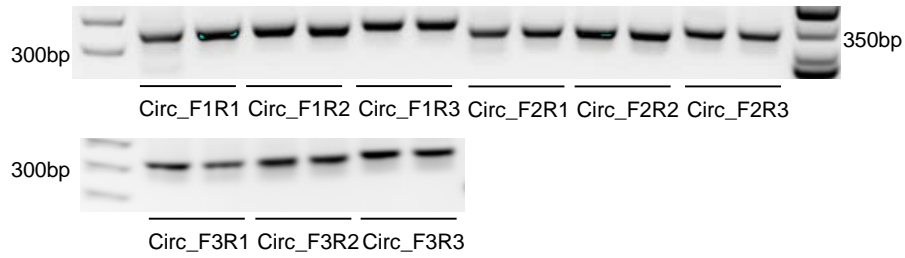
circTtn_chr2:76847388|76881839

circTtn_chr2:76847388|76884197
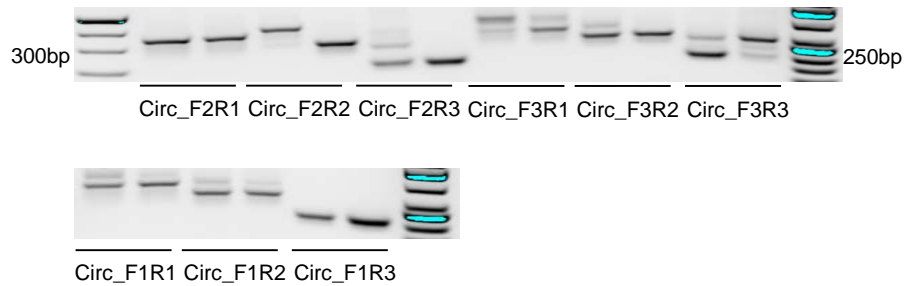
Figure 3-4. Electrophoresis of circRNAs in mouse genome after qPCR amplification. For each circRNAs, 3 forward primers and 3 reverse primers were designed and paired, resulting 9 primers set. 1 primer set for each of these 13 circRNAs was sent for Sanger sequencing and confirmed the existence.

circMyocd_chr10:51724607|51739115
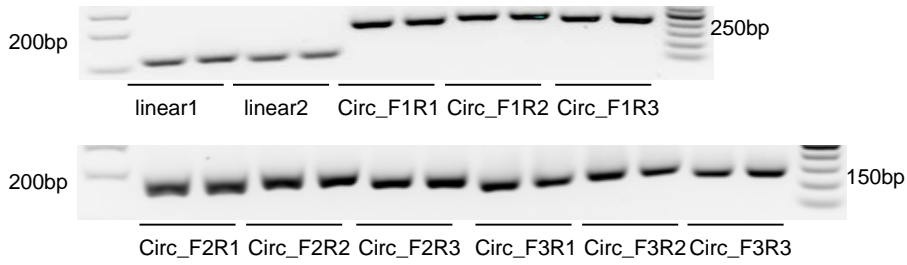
circZfp148_chr11:70584974|70570814
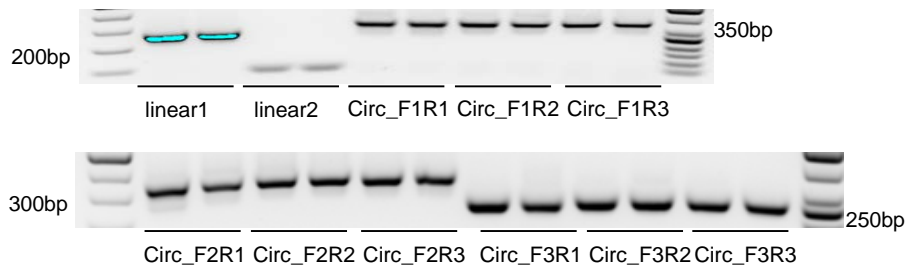
circRyr2_chr17:65735987|65781009
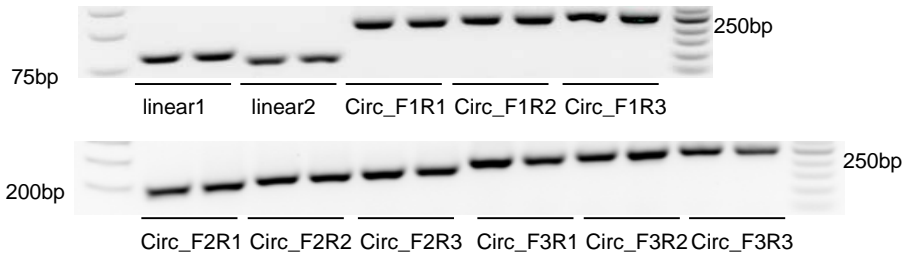
circRyr2_chr17:65750062|65781009
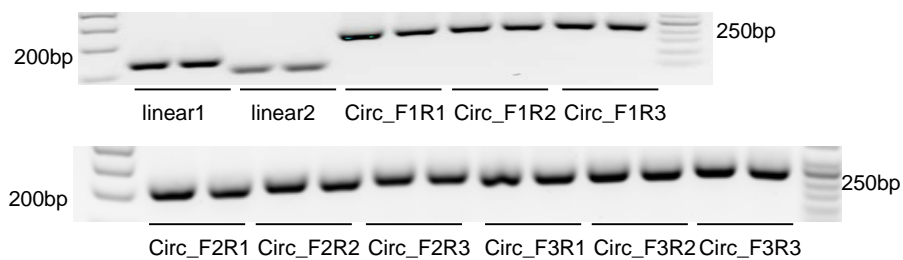
Figure 3-5. Electrophoresis of circRNAs in rat genome after qPCR amplification. For each circRNAs, 3 forward primers and 3 reverse primers were designed and paired, resulting 9 primers set. 1 primer set for each of these 4 circRNAs was sent for Sanger sequencing and confirmed the existence.

Figure 3-6. circMyocd and linear Myocd exhibited a negative correlation. A. Two siRNAs targeting the circMyocd back-spliced junction were designed. The red and green indicate the two exons that were spliced together. **B-C.** The siRNAs worked for circMyocd knockdown. **D-G**. Linear Myocd were checked using 2 different primers after circMyocd knockdowns.   * P <0.05

Table 3-1. Primers for circular and linear RNA confirmation in mouse genome

circMyocd_chr11:65218529|65233184	F1	CGGCAGAAAGGTCCATTCCA	R1	CCTGGGTCCTTCTCTGTTGA
	F2	CCACGGCAGAAAGGTCCAT	R2	GCGGTATTAAGCCTTGGTTAGC
	F3	GGTCCATTCCAACTGCTCAGA	R3	TTAAGCCTTGGTTAGCCAGC
linear Myocd for circMyocd_chr11:65218529|65233184	F1	CGGCAGAAAGGTCCATTCCA	R1	ACGGAAGAATCCATCGGCAG
	F2	CCACGGCAGAAAGGTCCAT	R2	TCACGGAAGAATCCATCGGC
circTtn_chr2:76857904|76881839	F1	AGGCTCCAAAGAAACCTGCT	R1	GTGTAGTCCCCGCTGTCTTC
	F2	CCAGAGGCTCCAAAGAAACCT	R2	CGCTGTCTTCCACACTGAGA
	F3	GCTCCAAAGAAACCTGCTCC	R3	CACCGAGTCCTCGAATGACA
linear Ttn for circTtn_chr2:76857904|76881839	F1	GGCGGGTCTCCAGAAATCAA	R1	GTGTAGTCCCCGCTGTCTTC
	F2	CGGATCGTGAAGCAGGATGA	R2	CGCTGTCTTCCACACTGAGA
circTtn_chr2:76857904|76884197	F1	AGGCTCCAAAGAAACCTGCT	R1	AAGAGTCCTTTCCAGCCACG
	F2	CCAGAGGCTCCAAAGAAACCT	R2	TTGGCTACTTTGAGCACCGT
	F3	GCTCCAAAGAAACCTGCTCC	R3	GTTGCTCGCGTAGCAGGTAT
linear Ttn for circTtn_chr2:76857904|76884197	F1	ACGGTGCTCAAAGTAGCCAA	R1	AAGAGTCCTTTCCAGCCACG
	F2	TAACCGTGAGATTCGCCCTG	R2	TTGGCTACTTTGAGCACCGT
circTtn_chr2:76863274|76868085	F1	GCCTTGGGAAGAGCCCTATG	R1	TTTGGTTGTTAAGTACAGTTCCGC
	F2	CGGAACCATACACAGAGCCA	R2	TTGGTTGTTAAGTACAGTTCCGC
	F3	ACGAGAGGAAGCAGTCTACG	R3	GTTAAGTACAGTTCCGCTGTGC
linear Ttn for circTtn_chr2:76863274|76868085	F1	GCCTTGGGAAGAGCCCTATG	R1	CCATTCCTCGTGGACTTGCT
	F2	CGGAACCATACACAGAGCCA	R2	TTCCCACTCTTCCTCCCCTT
circTtn_chr2:76802231|76818816	F1	ACCCCCGGTTGAGTTCACTA	R1	CACCACCTCTTCCTTTGGCT
	F2	TTGAGGACCAGACGGTCGAA	R2	AGCACCACCTCTTCCTTTGG
	F3	CCTCTTGAGGACCAGACGGT	R3	CGCTTTTCAGCACCACCTCT
circTtn_chr2:76847388|76881839	F1	TGCCAGAAGTGCCGAAGAAA	R1	GTGTAGTCCCCGCTGTCTTC
	F2	GCCAGAAGTGCCGAAGAAAC	R2	CGCTGTCTTCCACACTGAGA
	F3	CCGAAGAAACCTGTGCCTGA	R3	CACCGAGTCCTCGAATGACA
circTtn_chr2:76847388|76884197	F1	TGCCAGAAGTGCCGAAGAAA	R1	AAGAGTCCTTTCCAGCCACG
	F2	GCCAGAAGTGCCGAAGAAAC	R2	TTGGCTACTTTGAGCACCGT
	F3	CCGAAGAAACCTGTGCCTGA	R3	GATCGGTGCTGTTCCAGTGA
circRyr2_chr13:11759671|11785141	F1	AGGGCAATGAACACTACGGG	R1	GGGTTTCAAGCGACATCTGC
	F2	CAATGAACACTACGGGCGCT	R2	TGAGGGTTTCAAGCGACATCT
	F3	GCACCAGGGCAATGAACAC	R3	GGTTTCAAGCGACATCTGCAA
linear Ryr2 for circRyr2_chr13:11759671|11785141	F1	AGGGCAATGAACACTACGGG	R1	CAGGATCTCGCCGTTCAGAG
	F2	CAATGAACACTACGGGCGCT	R2	TCGCCGTTCAGAGTGAACAT

Table 3-1. Primers for circular and linear RNA confirmation in mouse genome-continued

circRyr2_chr13:11680966|11688013	F1	TCAAAAACCACCGCCTGTACT	R1	GGAACTTGAAGATGTCTTGGGC
	F2	ATTTCAAAAACCACCGCCTGT	R2	GGCTTTTTCTCTGTCCTTGGC
	F3	CCACCGCCTGTACTTCTTGT	R3	ACCACTGATCTGTAGGAACTTGA
linear Ryr2 for circRyr2_chr13:11680966|11688013	F1	TCAAAAACCACCGCCTGTACT	R1	TTGGATGCATGTCCTCCAGTG
	F2	ATTTCAAAAACCACCGCCTGT	R2	GGATGCATGTCCTCCAGTGC
circRyr2_chr13:11769852|11785141	F1	TCGTCTCGTCCCCTACACTC	R1	GGGTTTCAAGCGACATCTGC
	F2	AAGAAACCCTCGTCTCGTC	R2	TGAGGGTTTCAAGCGACATCT
	F3	CGTCCCCTACACTCTTCTGG	R3	GGTTTCAAGCGACATCTGCAA
linear Ryr2 for circRyr2_chr13:11769852|11785141	F1	TCGTCTCGTCCCCTACACTC	R1	CTTCCAGGTGGTAGCCGTAG
	F2	AAGAAACCCTCGTCTCGTC	R2	TGATCAGGAGCTTCCAGGTG
circAtp5c1_chr2:10058996|10064311	F1	CTTGACTTTCAACCGCACCC	R1	TGAGGGCAGCCACTTCATTC
	F2	GACCTTGACTTTCAACCGCAC	R2	GAGGGCAGCCACTTCATTCT
	F3	TTGACCTTGACTTTCAACCGC	R3	ATAGCACCACAAAGCCCTCT
linear Atp5c1 for circAtp5c1_chr2:10058996|10064311	F1	CTTGACTTTCAACCGCACCC	R1	CAGCAGCCCCAGAGATGATT
	F2	GACCTTGACTTTCAACCGCAC	R2	CAGCAGCCCCAGAGATGATTT
circPpp2r3a_chr9:101125023|101153874	F1	GAGCCATCTGACTGGGACAG	R1	ACTCCAAGATCTGTTGACTGTG
	F2	GCCCTGAGCCATCTGACTG	R2	CACTCCAAGATCTGTTGACTGTGT
	F3	TGACTGGGACAGATTTGCGG	R3	TCCAAGATCTGTTGACTGTGTAGAA
linear Ppp2r3afor circPpp2r3a_chr9:101125023|101153874	F1	GAGCCATCTGACTGGGACAG	R1	TTCTTGGAGCTGAACCTGGG
	F2	GCCCTGAGCCATCTGACTG	R2	GGGCAGACTCCTCAGTTACG
circZfp148_chr16:33421321|33434974	F1	CAGTCTTCCAGGGCAATGGT	R1	ATGTGTCAGGTATAAGCCCATCC
	F2	CGGCATAGACGAAATGCAG	R2	TGTGTCAGGTATAAGCCCATCC
	F3	CGAAGTTTGCCTCACCAGGA	R3	AAAATGTGTCAGGTATAAGCCCA
linear Zfp148for circZfp148_chr16:33421321|33434974	F1	CAGTCTTCCAGGGCAATGGT	R1	TGCATACTGCAGTCCTTGGG
	F2	CGGCATAGACGAAATGCAG	R2	TCGTCTGCAGCAAGGATCTC

Table 3-2. Primers for circular and linear RNA confirmation in rat genome

circMyocd_chr10:51724607	51739115	F1	CTCCACGGCAGAAAGATCCA	R1	TTGGTTAGCCAGCTGCTCC
	F2	ATGAAGCTGAAGAGAGCCCG	R2	GTATTAAGCCTTGGTTAGCCAGC
	F3	ACGAGAAGATTGCTCTCCGC	R3	CTGCTCCTGGGTCCTTCTCT
linear Myocd for circMyocd_chr10:51724607	51739115	F1	CTCCACGGCAGAAAGATCCA	R1	GCGGAGAGCAATCTTCTCGT
	F2	ATGAAGCTGAAGAGAGCCCG	R2	TTCTCCACCAGTTCCAAGGG
circZfp148_chr11:70584974	70570814	F1	CAGTCTTCCAGGGCAATGGT	R1	CCATCCCGCCAGATCACTTG
	F2	TGTGGCGGCATAGACGAAAT	R2	CATCCCGCCAGATCACTTGG
	F3	CGAAGTTTGCCTCACCAGGA	R3	CCGCCAGATCACTTGGCTA
linear Zfp148for circZfp148_chr11:70584974	70570814	F1	CAGTCTTCCAGGGCAATGGT	R1	AAGCGCATACTGCAGTCCTT
	F2	TGTGGCGGCATAGACGAAAT	R2	ACCATTGCCCTGGAAGACTG
circRyr2_chr17:65735987	65781009	F1	TGAACACTACGGGCGTTCTT	R1	GTAGTTGCGTTCCTGTTCCG
	F2	GTGGGCTGCATGGTTGATATG	R2	GCAAATTGTAGTTGCGTTCCTG
	F3	AATGAACACTACGGGCGTTCT	R3	CTGCAAATTGTAGTTGCGTTCC
linear Ryr2 for circRyr2_chr17:65735987	65781009	F1	TGAACACTACGGGCGTTCTT	R1	CATCCAGCAGGATCTCACCG
	F2	GTGGGCTGCATGGTTGATATG	R2	CAGCTCGGAGCCTGAGTCAT
circRyr2_chr17:65750062	65781009	F1	ACAGGAGAAACCCTCGTCTG	R1	GTAGTTGCGTTCCTGTTCCG
	F2	GGAGAAACCCTCGTCTGGTTC	R2	GCAAATTGTAGTTGCGTTCCTG
	F3	TCGTCTGGTTCCCTACACTC	R3	CTGCAAATTGTAGTTGCGTTCC
linear Ryr2 for circRyr2_chr17:65750062	65781009	F1	ACAGGAGAAACCCTCGTCTG	R1	GCTTCCAGGTGGTAGCCGTA
	F2	GGAGAAACCCTCGTCTGGTTC	R2	CTCGGAGACTGTCTTTGTTGGA

Table 3-2. Primers for circular and linear RNA in rat genome. Rat circMyocd (chr10:51724607|51739115) is the counterpart of mouse circMyocd (chr11:65218529|65233184) ; Rat circZfp148 (chr11:70584974|670570814) is the counterpart of mouse circZfp148 (chr16:33421321|33434974) ; Rat circRyr2 (chr17:65735987|65781009) is the counterpart of mouse circRyr2 (chr13:11759671|11785141) ; Rat circRyr2 (chr17:65750062|65781009) is the counterpart of mouse circRyr2 (chr13:11769852|11785141) .

References

1.   You X, Vlatkovic I, Babic A, Will T, Epstein I, Tushev G, Akbalik G, Wang M, Glock C, Quedenau C, Wang X, Hou J, Liu H, Sun W, Sambandan S, Chen T, Schuman EM, Chen W. Neural circular rnas are derived from synaptic genes and regulated by development and plasticity. *Nature neuroscience*. 2015;18:603-610

2.   Conn SJ, Pillman KA, Toubia J, Conn VM, Salmanidis M, Phillips CA, Roslan S, Schreiber AW, Gregory PA, Goodall GJ. The rna binding protein quaking regulates formation of circrnas. *Cell*. 2015;160:1125-1134

3.   Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, Loewer A, Ziebold U, Landthaler M, Kocks C, le Noble F, Rajewsky N. Circular rnas are a large class of animal rnas with regulatory potency. *Nature*. 2013;495:333-338

4.   Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S. Circrna biogenesis competes with pre-mrna splicing. *Molecular cell*. 2014;56:55-66

5.   Bahn JH, Zhang Q, Li F, Chan TM, Lin X, Kim Y, Wong DT, Xiao X. The landscape of microrna, piwi-interacting rna, and circular rna in human saliva. *Clinical chemistry*. 2015;61:221-230

6.   Burd CE, Jeck WR, Liu Y, Sanoff HK, Wang Z, Sharpless NE. Expression of linear and novel circular forms of an ink4/arf-associated non-coding rna correlates with atherosclerosis risk. *PLoS genetics*. 2010;6:e1001233

7.   Franklin S, Chen H, Mitchell-Jordan S, Ren S, Wang Y, Vondriska TM. Quantitative analysis of the chromatin proteome in disease reveals remodeling principles and identifies high mobility group protein b2 as a regulator of hypertrophic growth. *Molecular & cellular proteomics : MCP*. 2012;11:M111 014258

8.   O'Connell TD, Rodrigo MC, Simpson PC. Isolation and culture of adult mouse cardiac myocytes. *Methods in molecular biology*. 2007;357:271-296

9.   Rosa-Garrido M, Chapski DJ, Schmitt AD, Kimball TH, Karbassi E, Monte E, Balderas E, Pellegrini M, Shih TT, Soehalim E, Liem D, Ping P, Galjart NJ, Ren S, Wang Y, Ren B, Vondriska TM. High-resolution mapping of chromatin conformation in cardiac myocytes reveals structural remodeling of the epigenome in heart failure. *Circulation*. 2017;136:1613-1625

10. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. Star: Ultrafast universal rna-seq aligner. *Bioinformatics*. 2013;29:15-21

11. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. The sequence alignment/map format and samtools. *Bioinformatics*. 2009;25:2078-2079

12. Anders S, Pyl PT, Huber W. Htseq--a python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31:166-169

13. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome biology*. 2014;15:550

14. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*. 2005;102:15545-15550

15. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*. 2009;25:1754-1760

16. Gao Y, Wang J, Zhao F. Ciri: An efficient and unbiased algorithm for de novo circular rna identification. *Genome Biol*. 2015;16:4

17. Ueyama T, Kasahara H, Ishiwata T, Nie Q, Izumo S. Myocardin expression is regulated by nkx2.5, and its function is required for cardiomyogenesis. *Molecular and cellular biology*. 2003;23:9222-9232

18. Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide rnas mediate rna interference in cultured mammalian cells. *Nature*. 2001;411:494-498

19. Werfel S, Nothjunge S, Schwarzmayr T, Strom TM, Meitinger T, Engelhardt S. Characterization of circular rnas in human, mouse and rat hearts. *Journal of molecular and cellular cardiology*. 2016;98:103-107

20. Rybak-Wolf A, Stottmeister C, Glazar P, Jens M, Pino N, Giusti S, Hanan M, Behm M, Bartok O, Ashwal-Fluss R, Herzog M, Schreyer L, Papavasileiou P, Ivanov A, Ohman M, Refojo D, Kadener S, Rajewsky N. Circular rnas in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Molecular cell*. 2015;58:870-885

Chapter 4: Genome-wide DNA Methylation Analysis of Cardiac Surgery Patients

Abstract

DNA methylation is a stable, but selectively variable, epigenetic modification that can affect gene expression and chromatin structure and whose role in cardiovascular disease is not yet fully understood. Cardiovascular diseases account for 31% of all global deaths and many of these diseases require patients to undergo surgery. One of the more common side effects of cardiac surgery is the onset of post-operative atrial fibrillation (POAF). With approximately 30% of cardiac surgery patients receiving atrial fibrillation diagnoses, the pervasiveness of this condition is far-reaching. Complications arising from atrial fibrillation entail longer hospital stays, steeper medical expenses, lower rates of five-year survivorship, and an increased probability of dangerous comorbidities such as heart failure or stroke[1, 2]. We utilized reduced representation bisulfite sequencing (RRBS), a genome-wide DNA methylation sequencing method, to sequence the pre- and post-operative peripheral blood samples of 101 patients who underwent coronary artery bypass grafting (CABG) with or without valve replacement. RRBS selectively measures CpG sites around genes, or primary cytosine-phosphate-guanine sequences on the same strand, the genomic sites known to be particularly susceptible to methylation. We sought to establish the potential DNA methylation biomarkers by first determining pre- and post-operative CpG methylation variability, with the eventual goal of identifying methylation sites with POAF predictive properties. Through a paired t-test, we found 3,180 hypervariable CpG sites, 303 of which are located in promoter region and 2364 of which are located in the gene body region,

respectively. These CpG sites are associated with genes enriched in development, cellular adhesion and the immune and cardiovascular systems. Among the hypervariable CpG sites, transforming growth factor beta 1 (TGF-β1) and those in the integrin family were most promising because of their relationship with the immune and cardiac systems. Through the use of logistic regression analyses, certain CpG sites were found to potentially demonstrate the predictive ability of POAF including the CpG site associated with TGF-β1.

Introduction

DNA methylation is a stable and long-term epigenetic modification that operates not by altering the actual sequence of DNA itself, but by covalently transferring a methyl group (—CH3) to a nucleotide. Cytosine methylation--the most prevalent variant of methylation—is done by adding a methyl group to the fifth-positioned carbon of a cytosine ring, resulting in 5-methylcytosine (5-mC). This process modifies chromatin structure and gene expression. DNA methylation in promoter region has long been recognized as a gene silencing marker because of its role repelling and/or preventing transcriptional machinery from connecting to DNA[3]. However, the precise function of methylation in other genomic regions remains unclear, including methylation found in the gene body region that positively correlates with gene expression[4]. Unlike other epigenetic markers, methylation aids in the development of local heterochromatin[5].

Although generally stable, DNA methylation is not completely static, and changes in the methylome have implications for human diseases[6]. Recent studies show that DNA methylation alterations occur during the progression of various pathological heart

diseases and can be used as biomarkers[7-9]. Chen et al indicated that after treating two different mouse strains with isoproterenol, one of which was susceptible to isoproterenol and one of which was resistant to it, the differentially methylated fragments were close to genes involved in heart disease associated function for the sensitive strain while involved in cardiac muscle contractility for the resistant strain[8]. Meder et al demonstrated that there are distinct, differentially methylated CpGs conserved between left ventricular tissue and peripheral blood in patients suffering from dilated cardiomyopathy. These conserved CpGs are enriched mostly in cell adhesion[9]. However, these studies have relatively small sample sizes and were performed using control and experimental groups, rather than within the same patient. Furthermore, neither of these studies attempted to elucidate the DNA methylation adaptations that occur as a result of surgical treatments to heart disease, such as CABG or valve replacement. Some of these DNA methylation changes can be interpreted as potential biomarkers for POAF rendering improvements in treating the commensurate condition.

In this study, we used a human cohort of patients who underwent CABG with or without valve replacement. For each enrolled patient, we acquired clinical and demographic information, then utilized genome-wide DNA methylation sequencing on both pre- and post-operative peripheral blood samples. Hypervariable CpG sites and their associated genes were used to generate a post-operative landscape of DNA methylation changes at a single base pair resolution, with the purpose of isolating potential markers for atrial fibrillation.

Material and Methods

Patient enrollment

Our study was approved by the UCLA Institutional Review Board (IRB). We obtained written, informed consent from all participants for the usage of their clinical and demographic data, as well as molecular analyses of pre- and post-operative peripheral blood samples. The study is still ongoing and we are concurrently enrolling patients that undergo CABG with or without valve replacement at UCLA. Clinical data obtained includes information about age, sex, surgery specifications. The present of paroxysmal atrial fibrillation (PAF) and POAF were acquired from the electronic medical records. A survey was conducted to obtain patients' smoking histories and ethnicities. For this project, we analyzed a total of 101 patients' data (Table 4-1).

Peripheral blood processing

For each patient, two blood draws were collected, one pre-operative and the other one 24 hours post-operative. Genomic DNA was extracted from 200ul of blood using the PureLink Genomic DNA Mini Kit (Invitrogen, K1820-02) and according to the manufacturer's instructions. The purified Genomic DNA was stored at -20 °C.

Reduced representation bisulfite sequencing

The method was modified from the lab of Dr. Matteo Pellegrini[10]. Genomic DNA was digested with Mspl (NEB, R0106S) overnight in a 37 °C water bath. Digested DNA was purified by AMPure beads (Beckman Coulter, A63880). DNA end-repair and adenylation reaction contain the purified DNA, NEB buffer 2 (NEB, B7002S), Klenow fragment DNA exo (NEB, M0212S) and a dNTP mixture consisting of dGTP (NEB,

N0446S), dMeCTP (NEB, N0356S), dATP (NEB, N0446S). The reaction was then purified by AMPure beads. For the adapter ligation, TruSeq Illumina Adapters (Illumina, 15041757) were used with the purified DNA, quick ligation reaction buffer (NEB, M2200L) and quick ligase (NEB, M2200L). After AMPure purification and size selection, bisulfite conversion was performed with the EpiTech Fast Bisulfite Kit (Qiagen, 59824). PCR amplification was applied by MyTaq Mix (Bioline, Bio-25041) and Illumina Truseq PCR primer cocktail (Illumina, 15041757), followed by AMPure bead purification. DNA was then quantified using a Qubit fluorometer and run through 1% agarose gel. DNA fragments in the library were about 300-400bp. The DNA was then sent to the UCLA Genetics Core facility (lab of Dr. Matteo Pellegrini) for sequencing. The sequencing condition was 100bp, single-end and performed using Illumina HiSeq 2000 with 8 lanes/flow cell.

Alignment

The alignment procedure was modified from our lab's previous publication[8]. The reads were aligned to the human genome version hg38 and methylation calling was performed using BS-seeker2[11]. Since DNA methylation mostly occurs in the CpG context, we chose the CGmap file (details for C/G) generated after methylation call from BS-seeker2 and only focused on cytosine in CG dinucleotide context. We set the cytosine coverage cutoff at 10 to focus on more replicable fragments. The methylation level of the cytosine was calculated by the number of methylated cytosines to the total number of cytosines.

DNA methylation dataset

DNA methylation distribution was found using the R package Genomation[12] for different regions including: promoters, exons, introns, intergenic regions, CpG islands, CpG island shores and other region (excluding CpG island and CpG shore). Promoter region here is defined as 1.5kb upstream to 0.5kb downstream of the transcription start site. CpG island shore is defined as the region 2kb upstream and downstream of CpG island. The overall genome region is based on hg38 reference genome. All CpG is the region that selected from hg38 reference genome with the "CG" context. Measured CpG includes all the CpG measured in the study that covered by the RRBS libraries. In the event of regions overlapping among promoter, exonic, intronic and intergenic regions, the priority is defined as promoter > exon > intron > intergenic region.

Analyses of surgical differences versus genetics and environmental differences were performed in R on two randomly chosen patients (Patient A and Patient B). The p-value was obtained by a proportion test. Q-value was acquired by p-value adjustments made in R for multiple comparisons. Q-value for significance was set at 0.01.

Differential methylation analysis

Differential methylation analysis was performed using the RnBeads[13] package in R. For the preprocessing section, the CGmap files for 101 patients were first converted into "bismarkCov" format and imported into RnBeads. A total of 5,450,449 methylation sites were mapped to the hg38 reference genome. 4,204,611 methylation sites were retained after merging methylation information for both strands for each CpG. We filtered 28,566 SNP-enriched sites, 215 high coverage outlier sites and 3,739,781 sites

that do not present in all samples. 436,049 CpG sites were kept for the following analysis. Principal component analysis (PCA) was taken for all retained CpG sites (436,049 CpG sites), the retained CpG sites in the promoter region and gene body region.

A paired t-test was performed for the remaining 436,049 CpG sites. Hypervariable CpG sites were defined as those presenting a difference of the average methylation level larger than 10% between the pre- and post-operative states across all patients, with p-value being less than 10^{-5}. Heatmaps for the hypervariable CpG sites in promoter and gene body regions were generated using the pheatmap package in R with added annotations of the clinical and demographic information including: age, sex, surgery specification, PAF, POAF, smoking history, and ethnicity. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway were applied to the hypervariable CpG sites—in either hypo- or hypermethylated states—using the R package GOstats[14] for both promoter and gene body regions. The CpG sites were visualized in the UCSC genome with a custom track by uploading a bedgraph format file.

An unpaired t-test was applied for hypervariable CpG sites based on various demographical categories, including: age (≥ 60 years or < 60 years), sex, PAF, and valve surgery condition. P-value cutoff was set at 0.05. GO and KEGG pathway analyses were applied for the significant CpG sites as described above. We excluded: smoking history because there was not enough data; PAF since the data was asymmetric; and ethnicity because of the data dispersion.

Prediction analysis

Logistic regression was selected for the prediction of POAF in R. Age, sex, PAF condition and valve surgery condition were used as variables. Then, hypervariable CpG sites were implemented into the logistic model one-by-one with the four variables described. GO and KEGG pathway analysis for the significant CpG sites were performed. Additionally, penalized logistic regression model was built for all hypervariable CpG sites in promoter and gene body regions separately. P-value cutoff was set at 0.05.

Results

DNA methylation of CpG sites measured by RRBS

Measurement of genome-wide DNA methylation with single-based resolution was performed by RRBS using the methylation-insensitive restriction enzyme MspI. MspI digests the genomic DNA by recognizing and cutting C^CGG sites, resulting in the enrichment of genome areas containing a high CpG content (Figure 4-1). When comparing the CpG composition of promoter, exon, intron, and intergenic regions (Figure 4-1A-C), there was a clear increasing trend of CpGs in the promoter and exon regions that is consistent with previous literature[15]. Furthermore, our RRBS libraries were highly enriched for CpG islands and CpG island shores both of which influence transcription[16], providing an insightful landscape for dinucleotide CpG measurement (Figure 4-1D-F). Overall, the sequencing data was a strong source for investigating the DNA methylation of CpGs and related genes.

In terms of DNA methylation levels at global scale, our data exhibited a bimodal distribution for all CpG sites measured, and retained selected CpG sites after filtering for

defined parameters (Figure 4-2). These results were consistent with previous human and mouse studies using microarray and RRBS[8, 17, 18], confirming the practical viability of our dataset.

Genetics and environmental differences are stronger than surgical differences in human epigenomes

Two patients (patient A and patient B) were randomly selected from our 101-patient cohorts. For each CpG site measured in pre- and post-operative peripheral blood samples within both patients A and B, a proportional comparison test was performed with p-value adjustments for surgical variances, here considered an surgical differences. Patient genotypic and environmental differences between patients A and B were calculated for pre-operative peripheral blood samples using above method. The percentage of CpG sites with a significant q-value <0.01 was 0.2673% for genetics and environmental differences, comparing pre-operative samples of patients A and B. While for surgical differences, the percentages were 0.0019% (the pre- and post-operative samples of patient A) and 0.0001% (the pre- and post-operative samples of patient B). Apparently, there was a greater percentage of CpG sites showing significant variation for genetics and environmental differences than there were for surgical ones (Figure 4-3). These results indicate that both surgical, genetics and environmental changes influenced the human epigenome, with genetics and environmental variations accounting for a greater change. Therefore, it became pertinent to explore the potential effects of surgical changes, such as cardiac surgery, in our study on the landscape of DNA methylation.

Differential DNA methylation analysis as induced by cardiac surgery

To obtain a view of the DNA methylome at CpG sites, PCA analyses were applied to all retained CpG sites, the retained CpG sites in promoter and in gene body regions. From these scatterplots, a clear division of pre- and post-operative samples is shown for all three scenarios, indicating DNA methylation changes within cardiac surgery (Figure 4-4). At the same time, a visualization of any CpG site's methylation level difference between the pre- and post-operative conditions could be uploaded to the UCSC genome browser as a custom track, providing an interactive view window (Figure 4-5).

Of the 436,049 retained CpG sites, a total of 3,180 CpG sites were considered hypervariable (Figure 4-6). 303 of the hypervariable CpG sites were in promoter region whereas 2,364 of them were in gene body region. Of the 303 CpG sites in promoter region, 186 were hypo-methylated and the remaining 117 CpG sites were hyper-methylated in the post-operative state. Of the 2,364 hypervariable CpG sites in the gene body region, 1,620 of them were hypo-methylated post-operatively and 744 CpG sites were hyper-methylated.

Hypervariable CpG sites are mostly enriched in the immune system

In order to elucidate the potential pathways and biological processes influenced by these hypervariable CpG sites, GO and KEGG pathway analyses were performed. The following four groups were treated as inputs: genes associated with the hypo-methylated CpG sites in the promoter region; genes associated with the hyper-

methylated CpG sites in the promoter region; genes associated with the hypo-methylated CpG sites in the gene body region; and genes associated with the hyper-methylated CpG sites in the gene body region. For each group, the potential pathways and biological processes were generated by hypergeometric tests and selected for p-values < 0.05. Only the top 15 terms, as sorted by p-value, were visualized (Figure 4-7, Figure 4-8). Most of the terms in the GO and KEGG pathways were related to the immune system such as the terms related to T-cells or lymphocytes. Some of them were involved in cellular adhesion pathways already known to be involved in differentially methylated CpG sites of both ventricular tissue and blood[9]. Additionally, we observed matches for related metabolic and cardiac-related pathways such as hypertrophic cardiomyopathy (HCM) and arrhythmogenic right ventricular cardiomyopathy (ARVC).

Hypervariable CpG sites for patient group separation

Heatmaps were drawn to understand the distribution of hypervariable CpG sites for all 101 patients in promoter and gene body regions. According to Figure 4-9, the patients may be subdivided into different groups based on the different methylation patterns depicted. Annotations for clinical and demographic data were added to the heatmaps including age, sex, ethnicity, smoker history, valve surgery condition and POAF condition. We sought to define a connection between the demographic variables and the observed changes in methylation. However, no significant correlation was determined. Potentially, the compounding of difference demographic variables was

cause for methylation for methylation change variability at a level higher than any one variable itself.

We next analyzed CpG hypervariability with relation to select demographical data listed above, here being age, sex, POAF status and surgery type. T-tests were applied to each of these four groups to compare the differences in hypervariable CpG sites in promoter and gene body regions, respectively. The significant p-value was set at 0.05. GO and KEGG pathway analyses were subsequently performed for the genes associated with the significant CpG sites in the promoter and gene body regions. For the group based on age, 30 hypervariable CpG sites in promoter region and 280 hypervariable CpG sites in the gene body region were found to be significant. The pathways and biological process involved in the associated genes were mostly enriched in the immune system, cell adhesion and development (Figure 4-10). For the second group divided by POAF, 27 hypervariable CpG sites in the gene body region were significant and mostly enriched in cardiac-related and synthesis terms (Figure 4-11). For the groups separated based on sex, 3 hypervariable CpG sites in promoter region and 36 CpG sites in gene body region were chosen. The GO and KEGG pathway showed mostly metabolism and immune system biological process and pathways (Figure 4-12). The last group was divided by valve surgery condition. There were 30 and 259 significant hypervariable CpG sites in promoter and gene body regions respectively. Cardiovascular related and immune system pathways and biological process were enriched for the GO and KEGG pathway analysis (Figure 4-13).

TNF-beta and Integrin family might genes be good candidates for differentially methylation and patient group separation

Certain cardiovascular pathways from the previous results of GO and KEGG pathway analyses were highly distinguished in differential methylation and patient group separation analyses (Table 4-2). Thus, we looked into the input genes associated with significantly hypervariable CpGs in each analysis condition. TGF-β1 was involved in almost all of these cardiovascular queries, and it displayed a distinct difference when comparing the CpG site related to TGF-β1 for the pre- and post-operative states, in the valve surgery and POAF conditions (Figure 4-14). Connections were also identified between specific integrin family members include integrin subunit beta 7 (ITGB7), integrin subunit alpha 9 (ITGA9) and integrin subunit alpha 11 (ITGA11), and specific query terms, including hypertrophic cardiomyopathy (HCM), arrhythmogenic right ventricular cardiomyopathy (ARVC), and dilated cardiomyopathy. The significantly hypervariable CpG sites associated with TGF-β1 and integrin family genes could serve as potential the biomarkers for cardiovascular surgery.

Prediction of post-operative atrial fibrillation

Logistic regression was applied to measure the potential variables that may influence the incident of POAF. First, we input the clinical and demographic variables into the model. With the p-value cutoff set at 0.05, only age was shown to potentially affect POAF occurrence (Table 4-3).

Hypervariable CpG sites were then put into the regression model one-by-one with the clinical and demographic variables. Based on the p-value, 2 hypervariable CpG

sites in promoter region and 39 hypervariable CpG sites in gene body region were found significant including the CpG site associated with TGF-β1. Another CpG site that associated with PR domain zinc finger protein 1 (PRDM1) was found to be significant as well as in previous group separation for the valve surgery condition (Table 4-2). GO and KEGG pathway analyses were performed and the most significant results were related to development and cellular adhesion (Figure 4-15).

Discussion

In this study, we utilized a human cohort of 101 patients undergoing CABG with or without valve replacement surgery to identify potential biomarkers for POAF as a result of surgical stress. To achieve this, we used RRBS, a genome-wide methylation sequencing method, to analyze patient pre- and post-operative peripheral blood samples. Meanwhile, a t-test was conducted for patient group comparison in key demographics, including age, sex, POAF, and valve surgery condition. Additionally, we applied logistic regression analyses for POAF. The results indicated that most of hypervariable CpG sites were related to concern with the immune system, metabolism, cellular adhesion, and a myriad of cardiac-related biological processes. In our patient group analysis, we found that many of the significant hypervariable CpG sites pertaining to cardiovascular terms were enriched in POAF and valve surgery conditions. After looking into the cardiovascular term genes associated with significant CpG, we identified TGF-β1 and integrin family as being extensively variable.

TGF-β1 could be the bridge between the immune and cardiac systems. TGF-β1, a ubiquitously expressed pleiotropic cytokine and most prevalent TGF family member, is

associated with a wide variety of molecular functions including proliferation, cellular differentiation, fibrosis, cardiac remodeling and immune response. It can be secreted by most immune cells[19] and is represented in a wide variety of pathways such as IL-2-dependent T-lymphocyte differentiation and proliferation, B-cell differentiation, macrophage deactivation, decreased reactive oxygen generation, and monocyte chemotaxis[20]. TGF-β1 has long-known relations to atrial fibrillation and other cardiac-related diseases[20-24]. It acts as a direct downstream mediator of angiotensin II[25] and subsequently transduces fibrosis signaling pathways, a hallmark characteristic for atrial fibrillation[26]. Selective atrial fibrosis induced by the overexpression of active TGF-β1 in transgenic mouse hearts with normal ventricular function and structure exhibited increased chances of atrial fibrillation[24]. Meanwhile, Pirfenidone, an antifibrotic drug, reduced the incidence of atrial fibrillation in canine with a decreased level of TGF-β1 expression[27]. Thus, TGF-β1 may act as an intermediate for communication between cardiac and immune systems.

Integrins are a group of heterodimeric transmembrane receptor proteins that perform signal transduction involved in cell cycle regulation, proliferation, and differentiation, as well as to aid in the adhesion of cells to the extracellular matrix (ECM)[28]. In the immune system, integrins are cardinal molecules in facilitating adhesion[29] whereas in the cardiac system, integrins transmit hypertrophic signaling[30] and play a role cardiac fibrosis[31]. Interestingly, integrins can bind and promote the function of TGF-β1. Targeting integrins attenuated organ fibrosis by regulating TGF-β1[32-34], providing a potential mechanism for drug application. Further study opportunities

lie in elucidating the manners in which integrin family genes and TGF-β1 may work together to facilitate crosstalk between the immune and cardiovascular systems.

Meanwhile, logistic regression analyses were applied to the hypervariable CpG sites. When calculating the clinical and demographic variables for the predictive ability of POAF, only age was significant, indicating that age may influence the incidence of POAF. Meanwhile, each hypervariable CpG sites was input into the model as well as age, sex, valve surgery and PAF conditions. Gene enrichment calculations for the significant hypervariable CpG sites revealed a concentration for genes associated with immune system, development, and cellular adhesion. It was noteworthy that the CpG site associated with TGF-β1 was also in the significant CpG site list for the logistic regression, further demonstrating its potential predictive capacity. Another CpG site associated with PRDM1 may also have certain prediction power for POAF. PRDM1, also called B lymphocyte induced maturation protein 1 (BLIMP-1), is a transcription factor that represses the beta-interferon expression and mediates B-cell as well as T-cell differentiation[35, 36]. Additionally, PRDM1 contributes to the heart morphogenesis[37].

We also performed the penalized logistic regression with lasso and elastic net using all the hypervariable CpG sites in the promoter and gene body regions separately. The results did not show any significant CpG sites with the lambda calculated by cross-validation, indicating the need to increase patient number (data not shown).

A limitation of this study is the volume of patients enrolled, though we are actively recruiting more patients for the project. Another concern for the study is that part of the hypervariable CpG sites between the pre- and post-operative samples may cause by the cardiac surgeries. Due to the physical stress of surgery on the heart, the immune

system may response with cell type-specific alterations, ultimately impacting the DNA methylome, a phenomenon of great interest to our lab and one for which we are currently designing novel analyses. One solution for this is to look at the cell type-specific DNA methylation patterns for the prediction and this study is also undergoing in our lab.
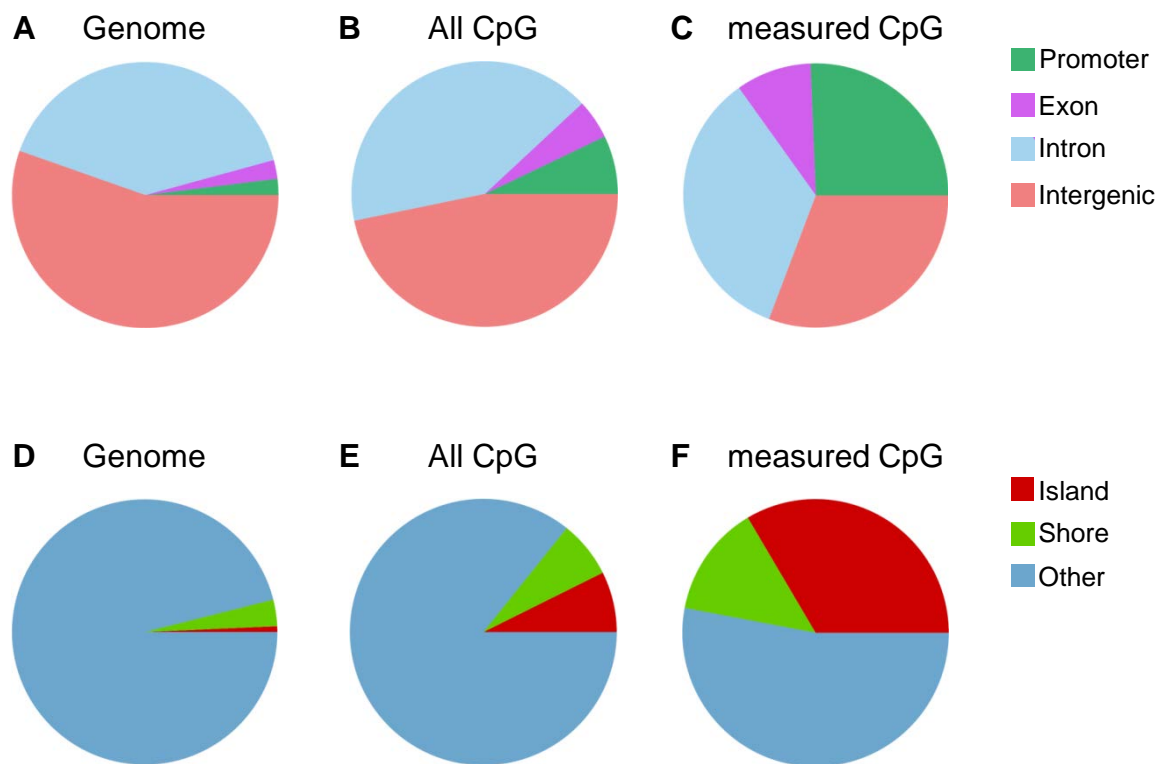
Figure 4-1. Global distribution of DNA methylation. A-C. CpG sites were categorized based on their positions in the genome namely promoter, exon, intron, and intergenic regions. There is an enrichment of the promoter and exon regions in RRBS. **D-F.** CpG sites were subdivided into CpG islands, CpG island shores, and other. The measurement had a high portion of CpG sites in the island and island shore regions. Promoters were here defined as 1.5kb upstream and 0.5kb downstream of the transcription start site. The CpG island shore was here defined 2kb upstream and downstream of the CpG island.
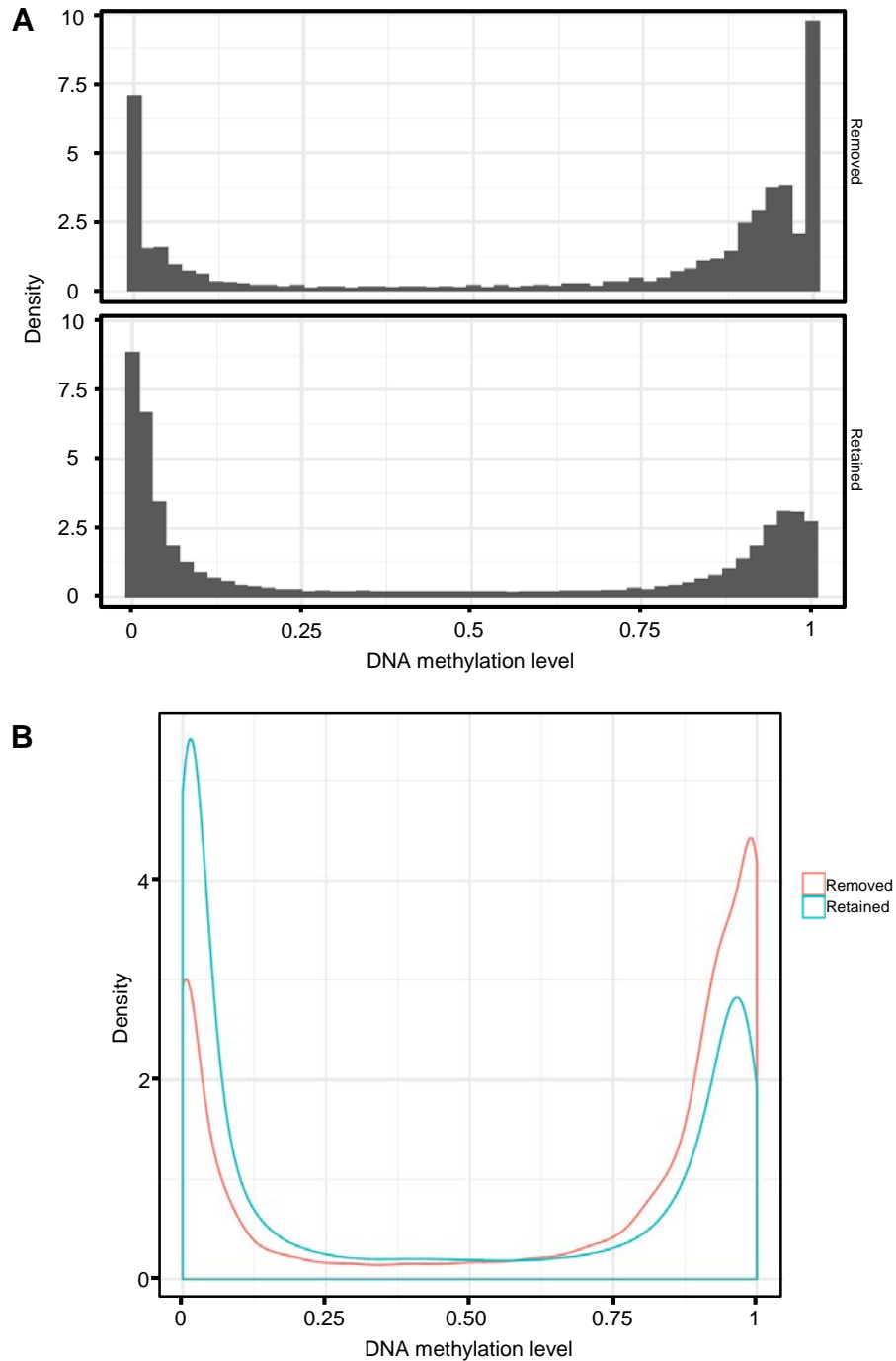
Figure 4-2. Distribution and comparison of the DNA methylation level for removed and retained CpG sites. A. Both removed and retained CpG sites exhibited a binomial distribution of DNA methylation level. **B.** When comparing the remove and retained CpG sites, the latter have a high portion of DNA methylation level around 0%, whereas the former have a high percentage in the DNA methylation of 100%.
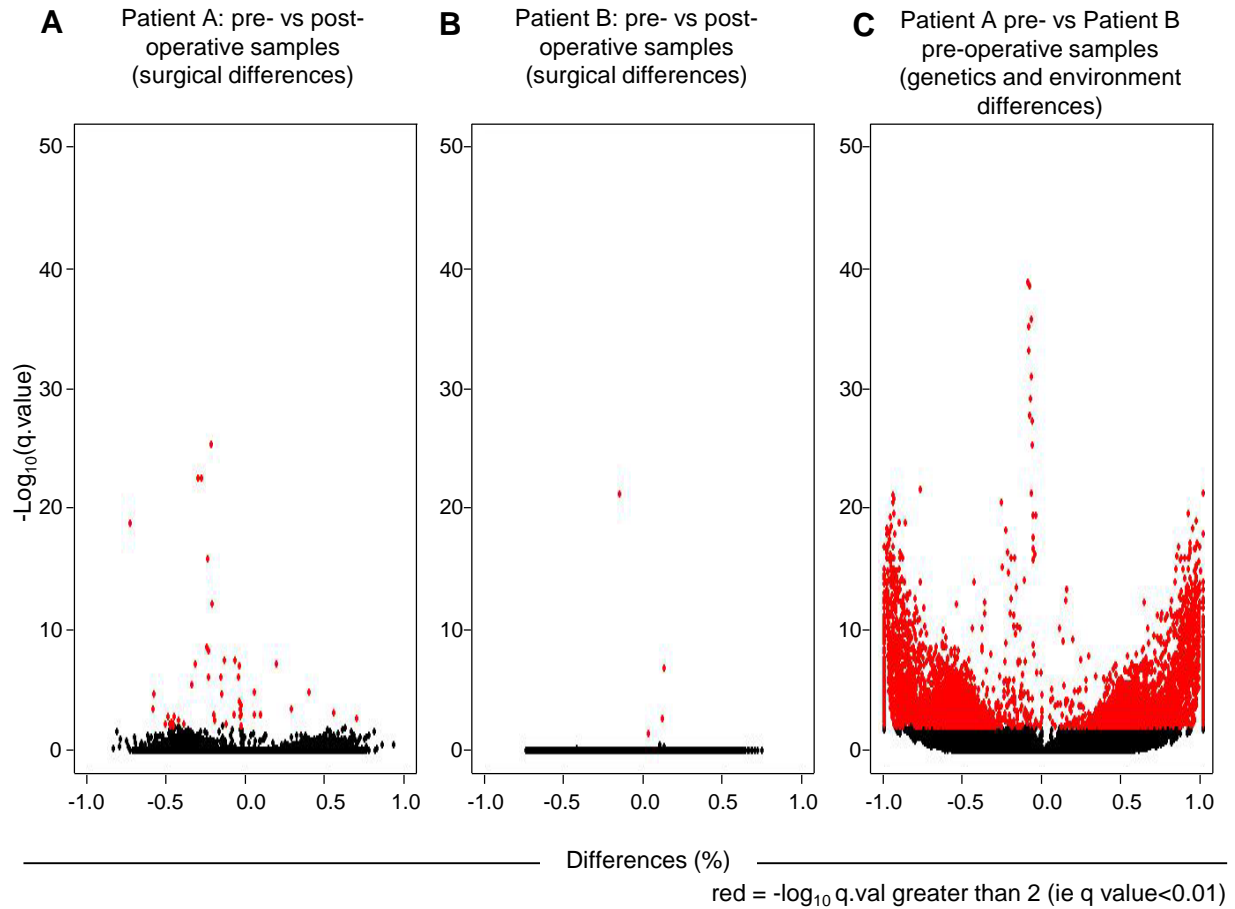
131

A Patient A: pre- vs post-operative samples (surgical differences)

B Patient B: pre- vs post-operative samples (surgical differences)

C Patient A pre- vs Patient B pre-operative samples (genetics and environment differences)

Differences (%)

red = $-\log_{10}$ q.val greater than 2 (ie q value<0.01)

Figure 4-3. Comparison of surgical, genetics and environmental differences in human epigenomes. A. Comparison between the pre- and post-operative samples' CpG sites for Patient A, representing methylation difference per surgical stimulus. **B.** Comparison between the pre- and post-operative samples' CpG sites for Patient B, representing methylation difference per surgical stimulus. **C.** Comparison between the CpG sites of patients A and B using pre-operative samples, indicating the genetics and environmental differences. There are more CpG sites demonstrating significant differences in genetics and environmental differences (C) than that of surgical differences (A-B), supporting the idea that genetics and environmental variability may be a stronger influence than surgical ones in human epigenomes. Q-value = 0.01.
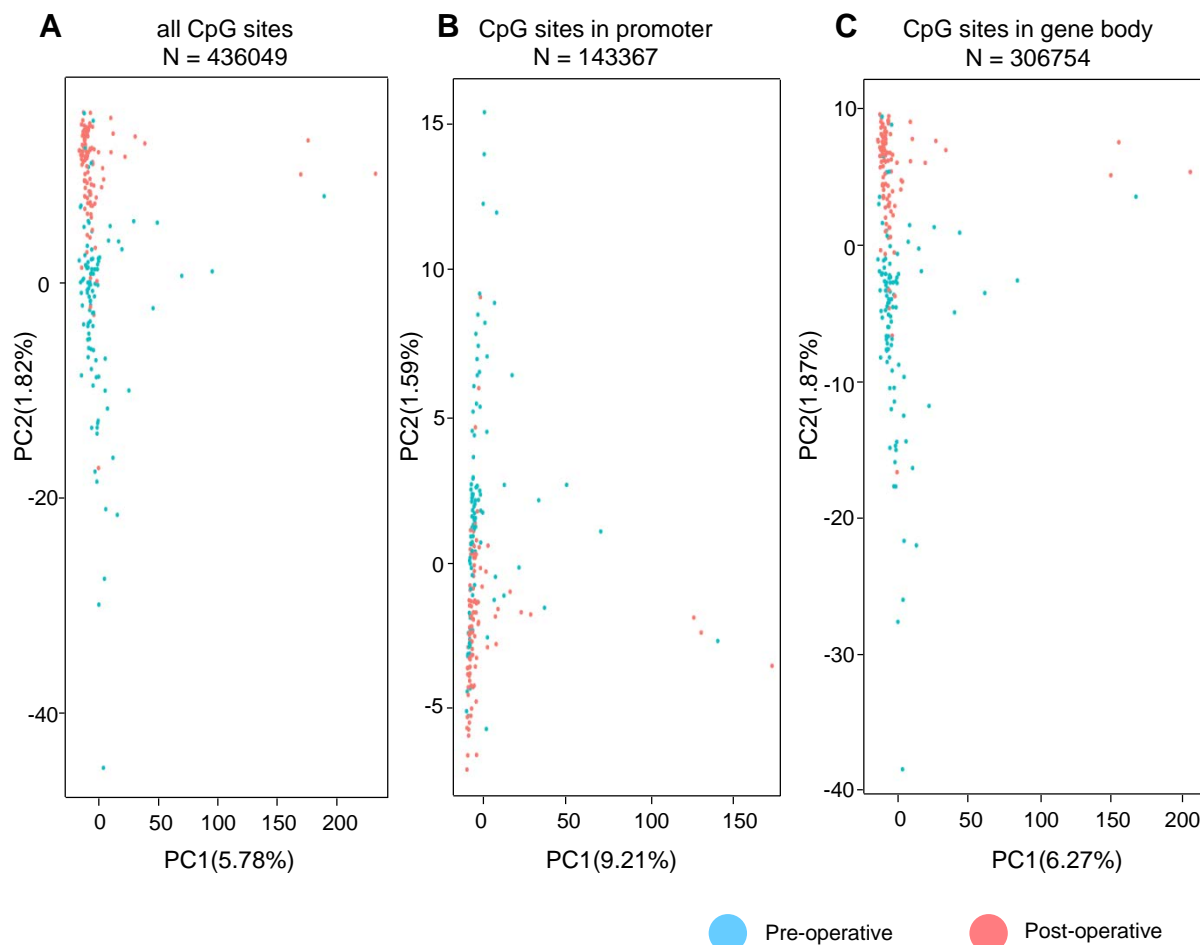
Figure 4-4. DNA methylation patterns for all CpG sites, those in promoter region, and those in gene body region. A. PCA for all retained CpG sites. **B.** PCA for retained CpG sites in the promoter region. **C.** PCA for retained CpG sites in the gene body region. All three above conditions showed distinct patterns for the pre- and post-operative CpG sites, suggesting there are CpG sites differences for the pre- and post-operative samples in both promoter and gene body regions. The blue dots indicate the pre-operative samples and the pink dots indicate the post-operative samples.

A Transforming growth factor beta 1 (TGF-β1)
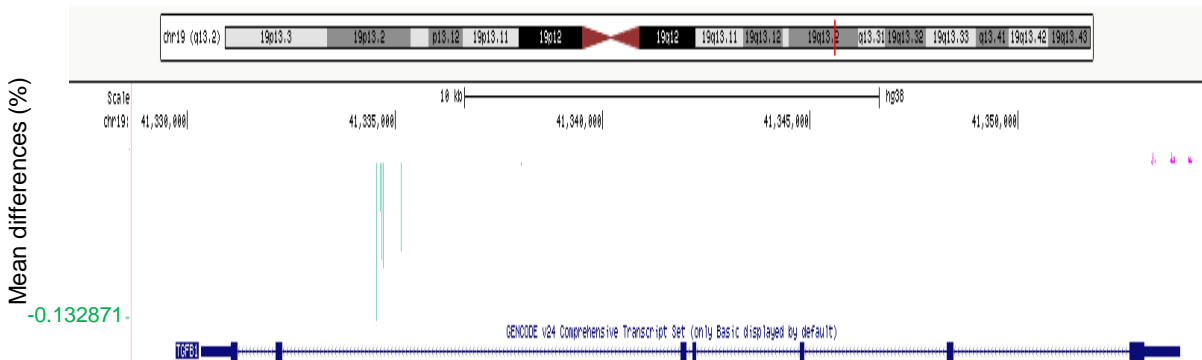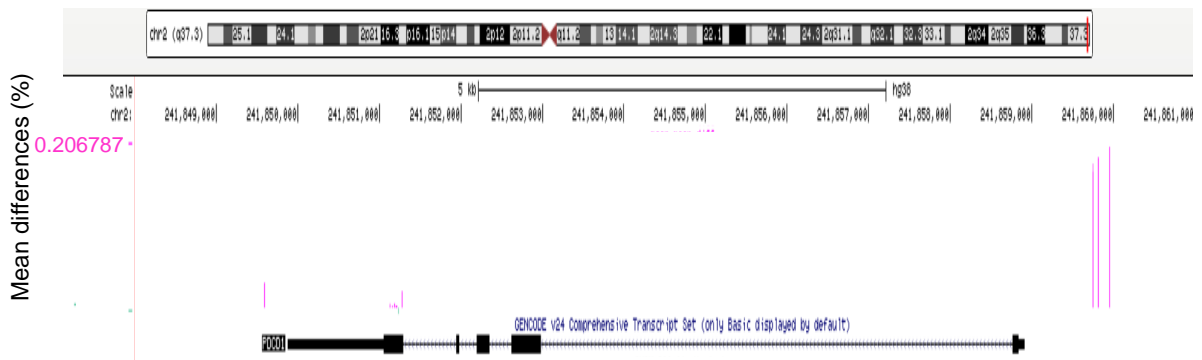

B Programmed cell death 1 (PDCD1)

Figure 4-5. Visualization of CpG sites for specific genes. A. Most of the CpG sites in the gene body region of transforming growth factor beta 1 (TGF-β1) were hypo-methylated. **B.** In the promoter region, programmed cell death 1 (PDCD1) was hyper-methylated. Each line (red or green) indicates a single CpG site. Green indicates the CpG site is hypo-methylated and red indicates the CpG site is hyper-methylated.
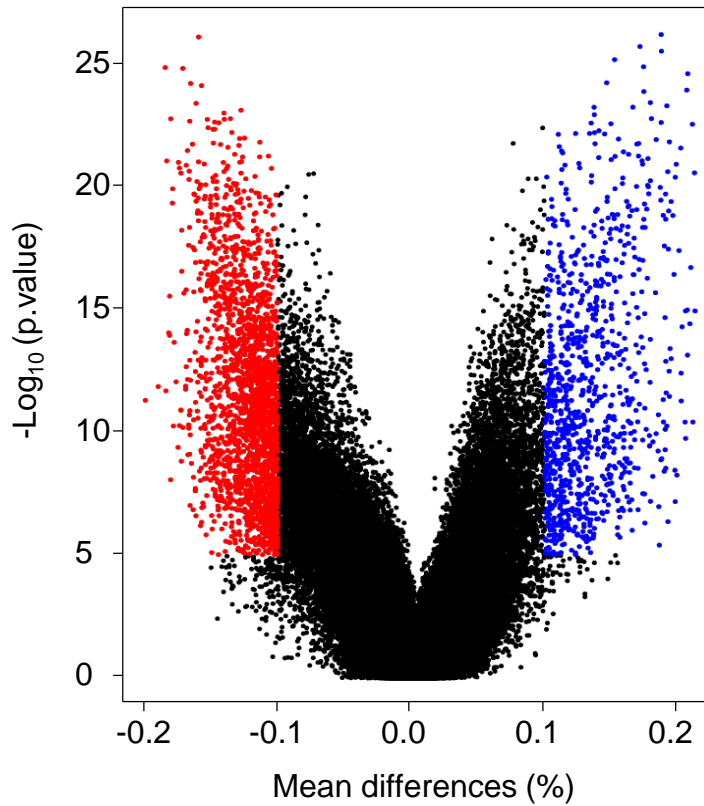
Figure 4-6. Volcano plot for DNA methylation difference of all CpG sites. Hypervariable CpG sites were identified via the difference between the pre- and post-operative samples being greater than 10% and the p-value from the paired t-test being less than 10^{-5}. The red dots indicate the hypervariable CpG sites that are hypo-methylated and the blue dots indicate the hypervariable CpG sites that are hyper-methylated.
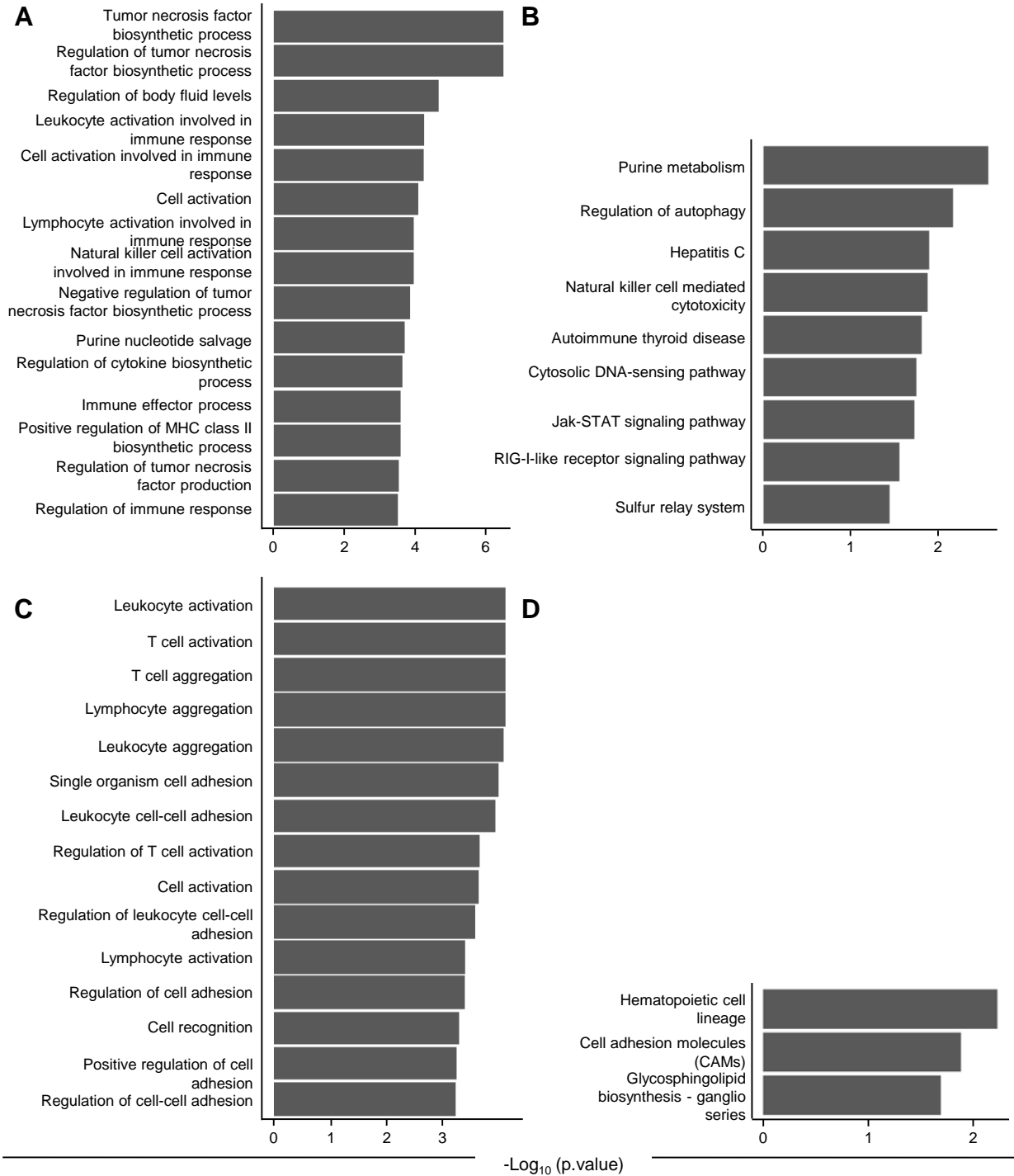
Figure 4-7. Gene enrichment analysis for hypervariable CpG sites associated genes in promoter region. A. GO analysis for associated genes of hypo-methylated CpG sites. **B.** KEGG pathway analysis for associated genes of hypo-methylated CpG sites. **C.** GO analysis for associated genes of hyper-methylated CpG sites. **D.** KEGG pathway analysis for associated genes of hyper-methylated CpG sites. Most of the enriched genes are in the immune system.
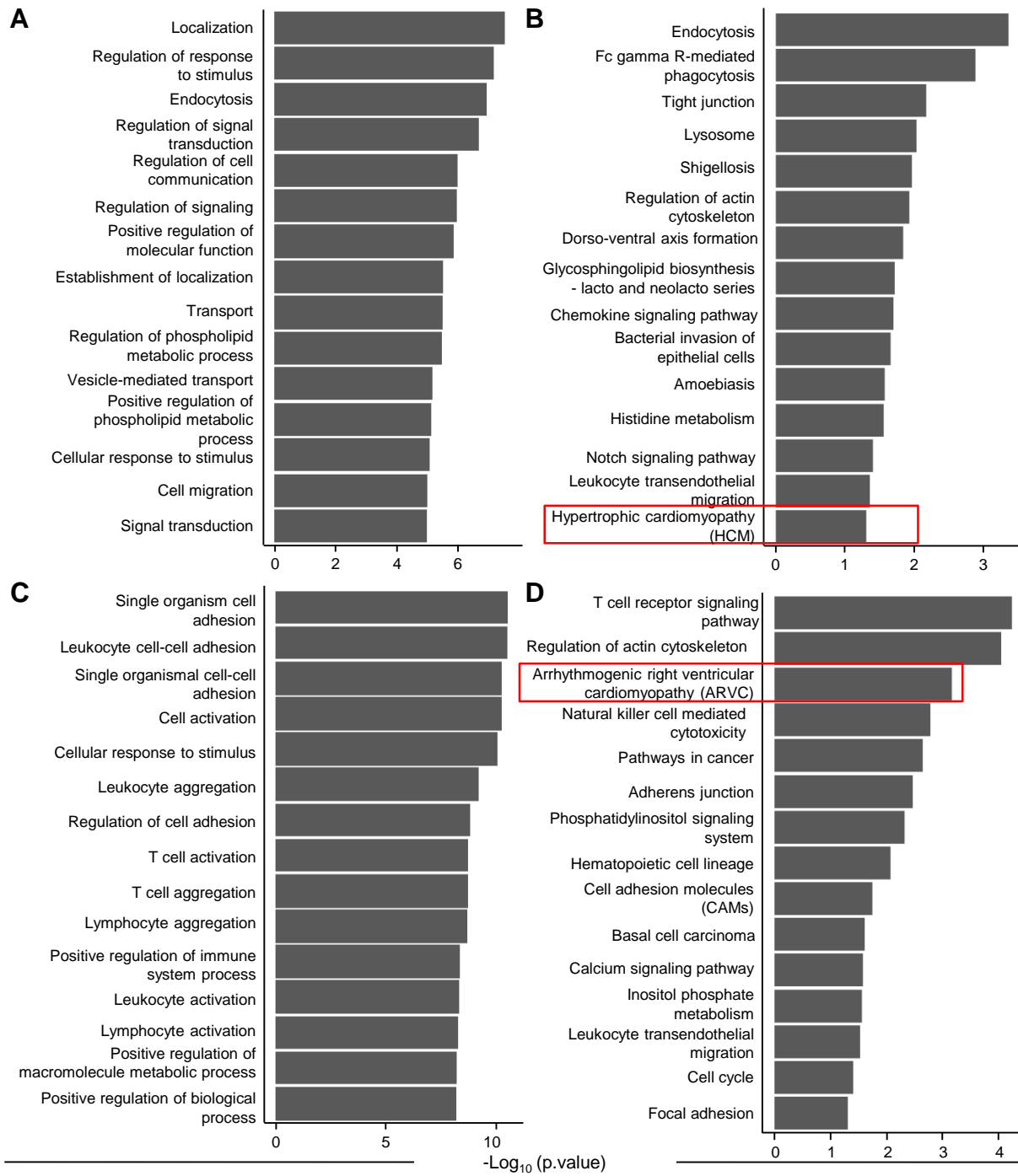
Figure 4-8. Gene enrichment analysis for hypervariable CpG sites associated genes in gene body region. A. GO analysis for associated genes of hypomethylated CpG sites. **B.** KEGG pathway analysis for associated genes of hypomethylated CpG sites. **C.** GO analysis for associated genes of hyper-methylated CpG sites. **D.** KEGG pathway analysis for associated genes of hyper-methylated CpG sites. Most of the enriched genes are in cellular adhesion, immune system, and cardiovascular system.
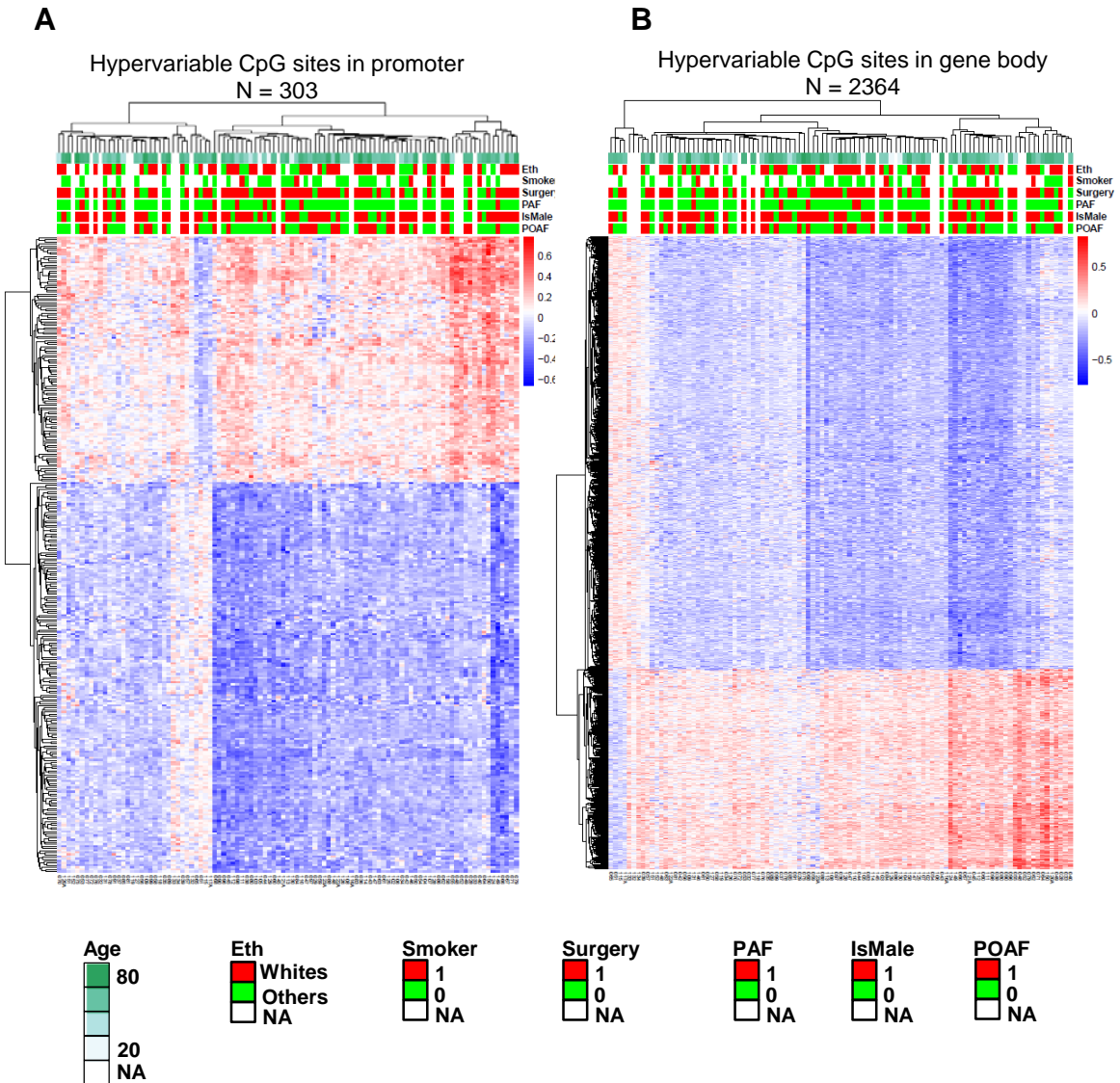
137

A

Hypervariable CpG sites in promoter
N = 303

B

Hypervariable CpG sites in gene body
N = 2364

Age	Eth	Smoker	Surgery	PAF	IsMale	POAF
80	Whites	1	1	1	1	1
	Others	0	0	0	0	0
20	NA	NA	NA	NA	NA	NA
NA						

Figure 4-9. Heatmaps with annotations for hypervariable CpG sites in promoter and gene body regions. A. Heatmap for hypervariable CpG sites in promoter region with annotations for age, ethnicity, smoking history, valve surgery condition, PAF, sex, and POAF. **B.** Heatmap for hypervariable CpG sites in gene body region with annotations of age, ethnicity, smoking history, valve surgery condition, PAF, sex, and POAF. The x-axis represents the patient number, and the y-axis represents the CpG site. The patients were divided into methylation difference groups, but none of the annotations alone could explain the group difference.
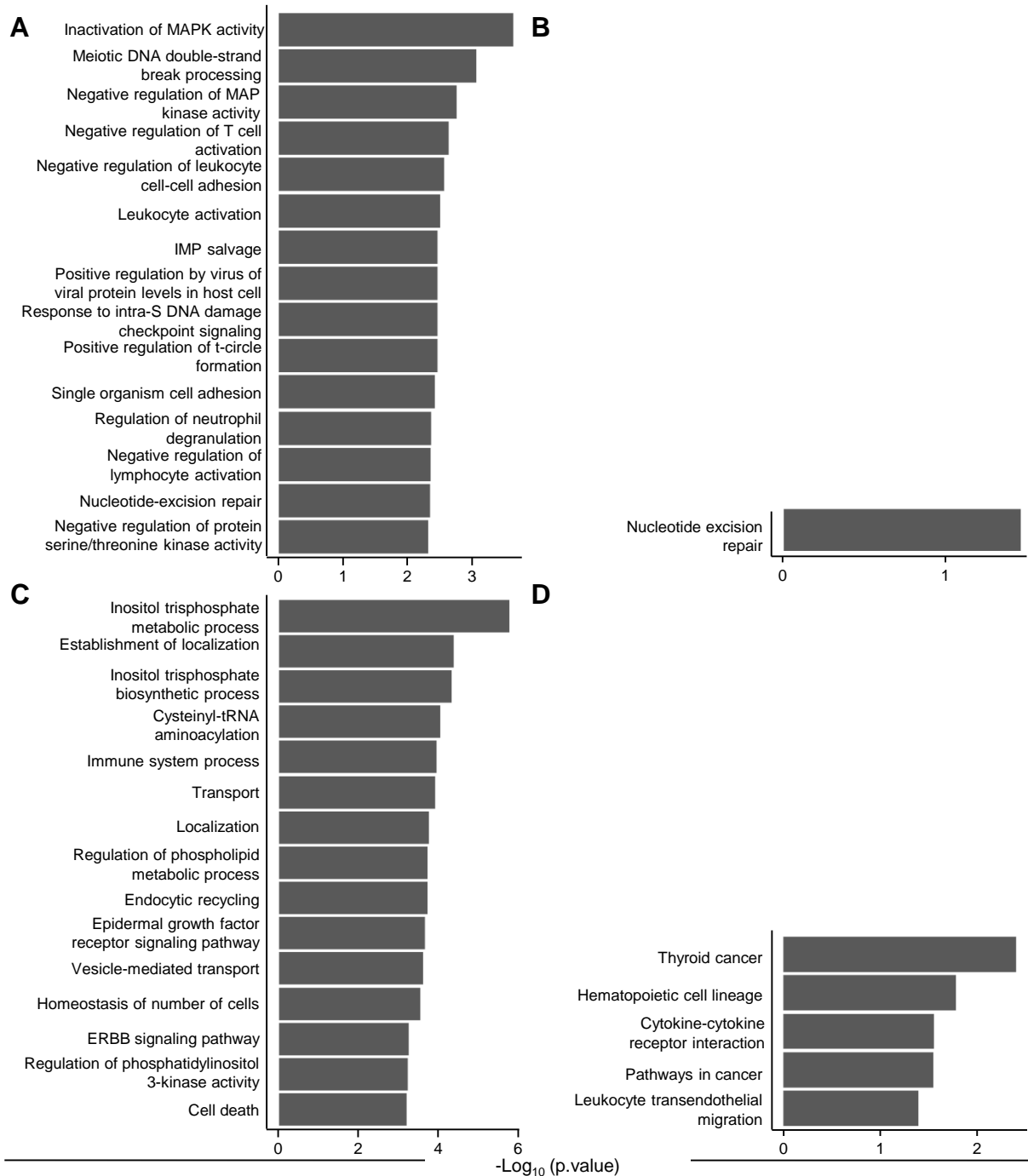
Figure 4-10. Gene enrichment analysis for associated genes of significant CpG sites in promoter and gene body regions when comparing age group ≥ 60 and age group < 60. A. GO analysis for associated genes of significant CpG sites in promoter region. **B.** KEGG pathway analysis for associated genes of significant CpG sites in promoter region. **C.** GO analysis for associated genes of significant CpG sites in gene body region. **D.** KEGG pathway analysis for associated genes of significant CpG sites in gene body region. Most of the gene enrichment are in immune system and metabolism.
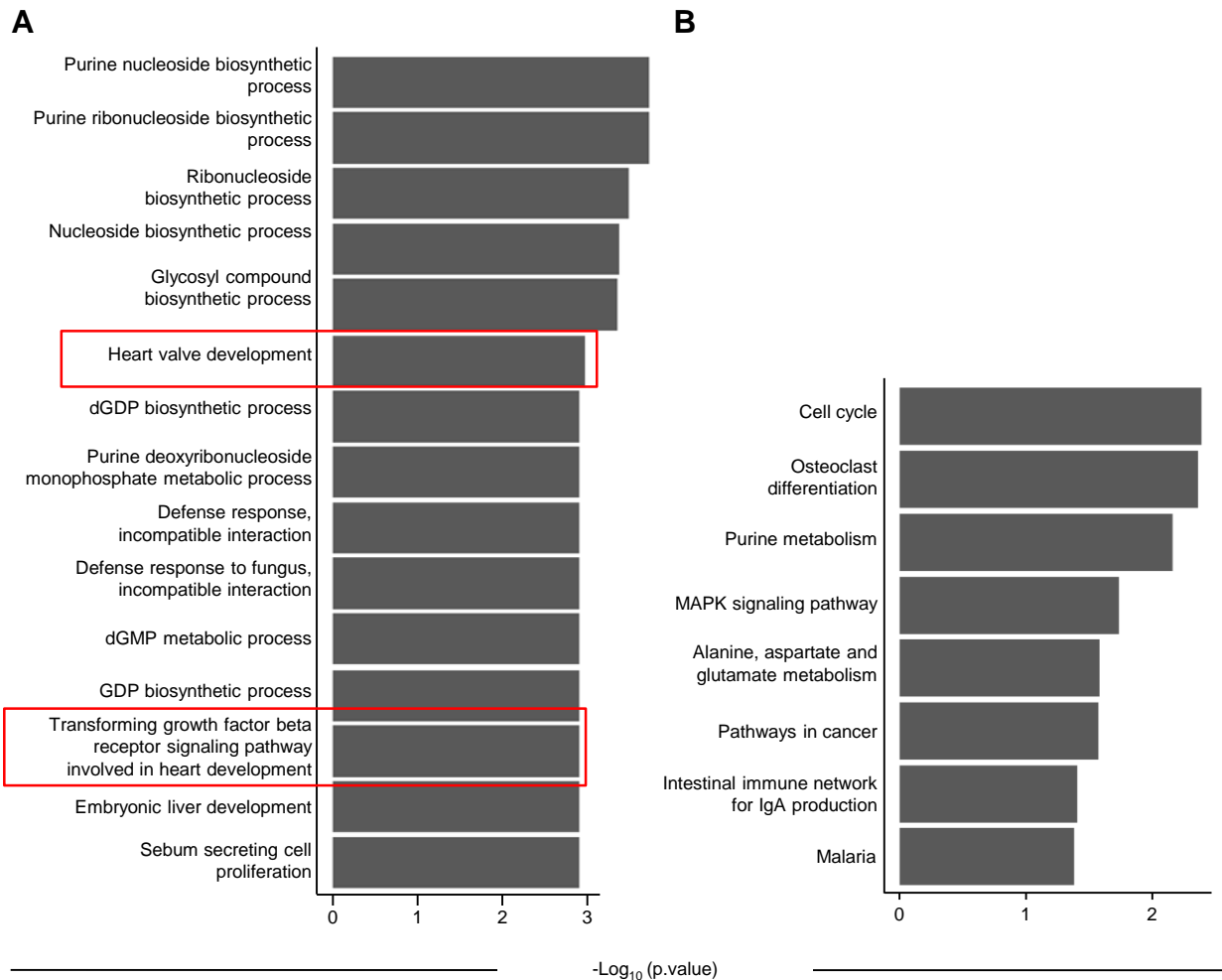
Figure 4-11. Gene enrichment analysis for associated genes of significant CpG sites in gene body region when comparing the POAF group and non-POAF group. A. GO analysis for associated genes of significant CpG sites in gene body region. **B.** KEGG pathway analysis for associated genes of significant CpG sites associated genes in gene body region. Most of the enriched genes are in biosynthesis, metabolism, and cardiac terms.
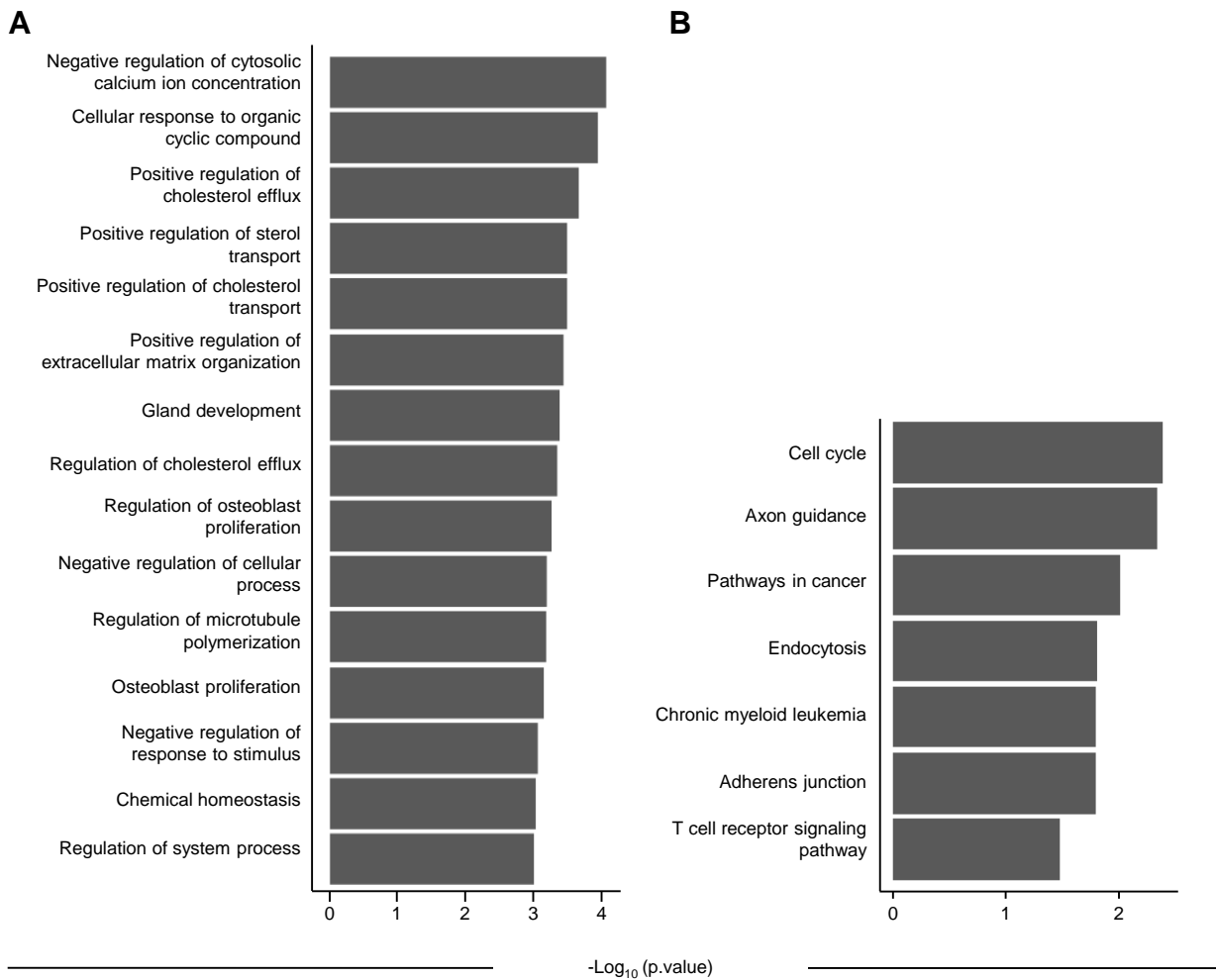
Figure 4-12. Gene enrichment analysis for associated genes of significant CpG sites in gene body region when comparing male with female. A. GO analysis for associated genes of significant CpG sites in gene body region. **B.** KEGG pathway analysis for associated genes of significant CpG sites in gene body region. Most of the enriched genes are in cell cycle, immune system, and metabolism.
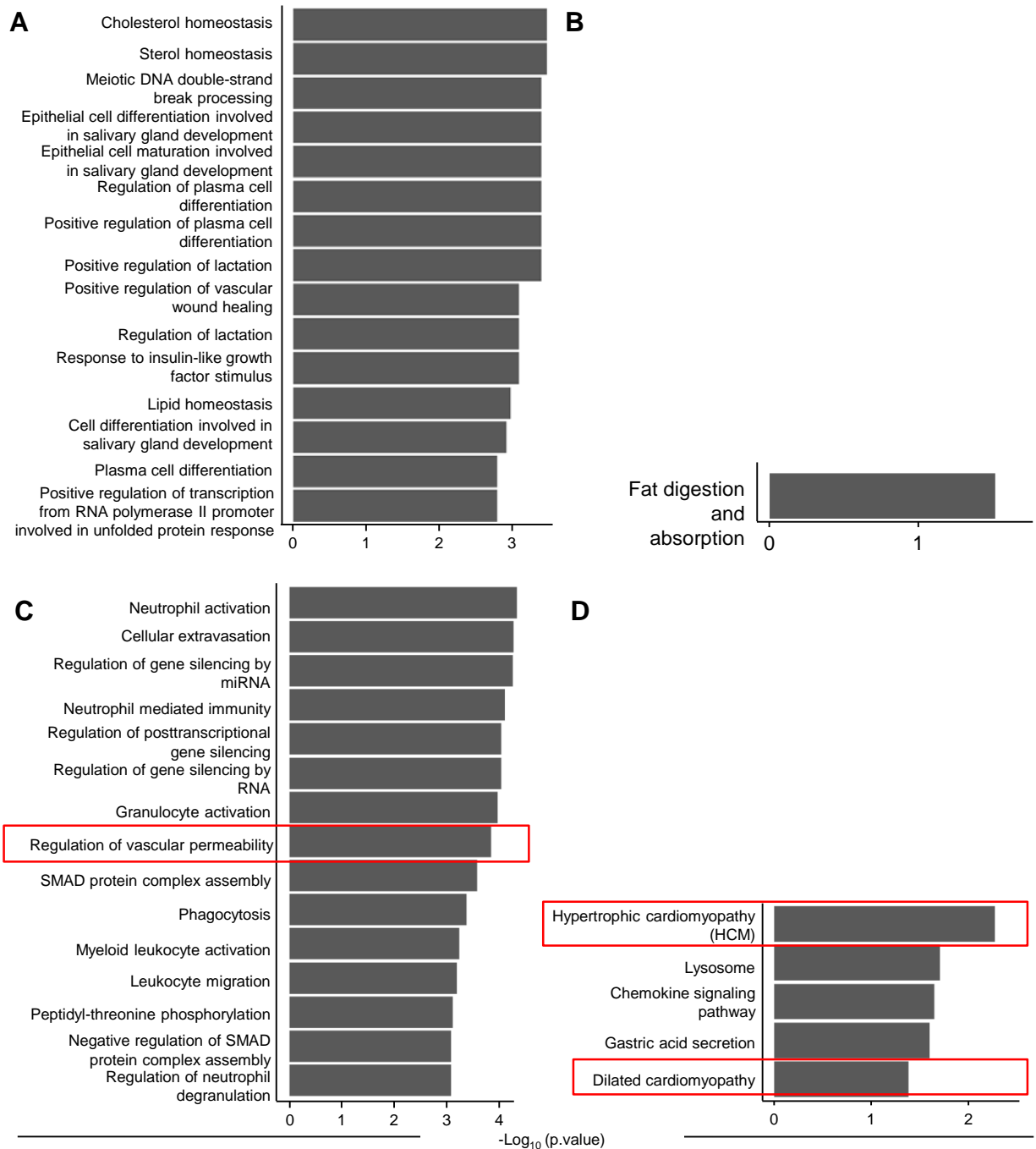
Figure 4-13. Gene enrichment analysis associated genes of significant CpG sites in promoter and gene body regions when comparing group with valve surgery and group without valve surgery. A. GO analysis for associated genes of significant CpG sites in promoter region. **B.** KEGG pathway analysis for associated genes of significant CpG sites in promoter region. **C.** GO analysis for associated genes of significant CpG sites in gene body region. **D.** KEGG pathway analysis for associated genes of significant CpG sites in gene body region. Most of the enriched genes are in cardiovascular system, metabolism, and immune system genes.
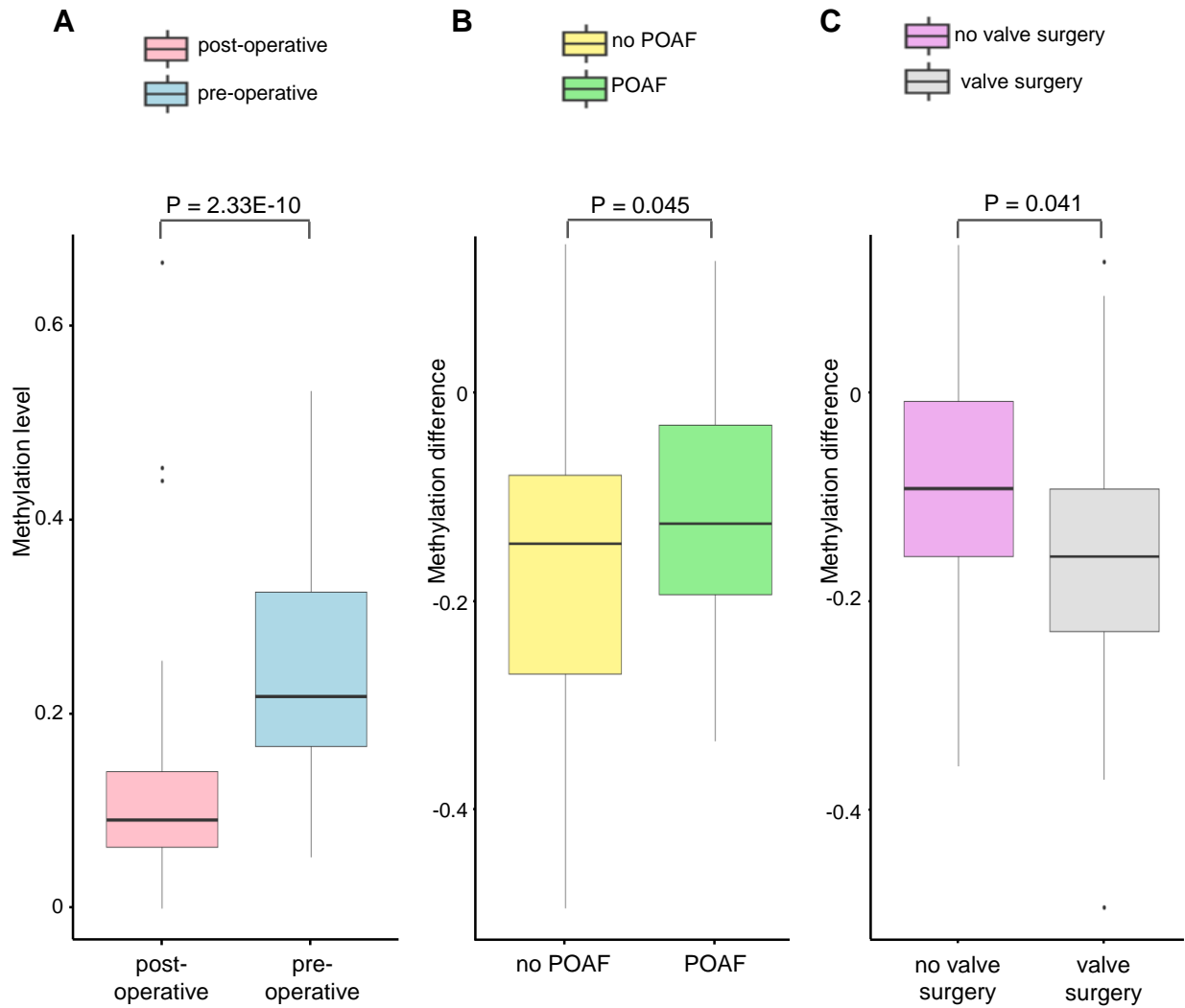
Figure 4-14. DNA methylation level comparison of the CpG site in TGF-β1. A. Comparing DNA methylation level of the CpG site in TGF-β1 for the pre-operative and post-operative groups. **B.** Comparing DNA methylation difference levels of the CpG site in TGF-β1 for the group with POAF and the group without POAF. **C.** Comparing DNA methylation difference level of the CpG site in TGF-β1 for the group with valve surgery and the group without valve surgery. All three comparisons revealed distinct DNA methylation differences for TGF-β1.

143

Figure 4-15. Gene enrichment analysis for associated genes of significant CpG sites for POAF prediction in gene body region. A. GO analysis for associated genes of significant CpG sites in gene body region. **B.** KEGG pathway analysis for associated genes of significant CpG sites in gene body region. Most of the enriched genes are in development, immune system, and metabolism.
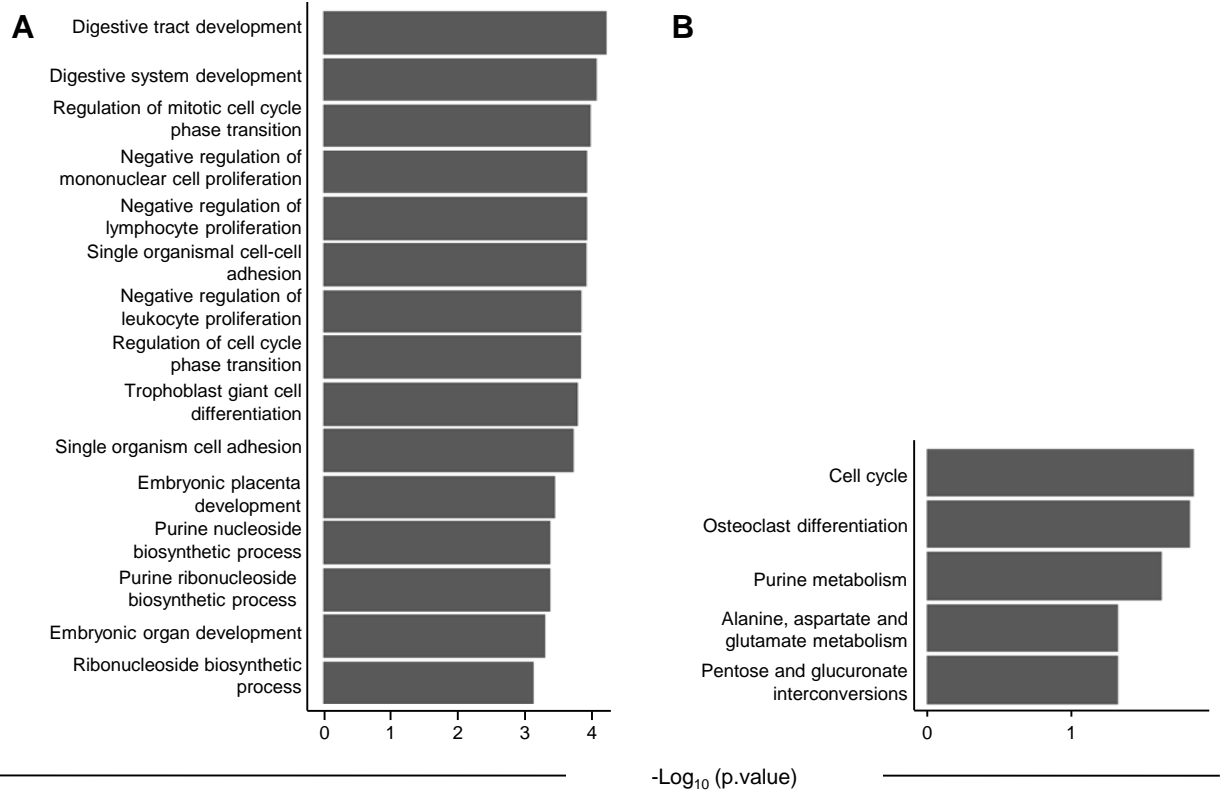
Table 4-1 Demographic and clinical characteristic of 101 enroll patients that underwent CABG

A

	Yes	No	Unknown	Total
POAF*	35 (34.6%)	48 (47.5%)	18 (17.8%)	101 (100%)
Male	60 (59.4%)	23 (22.8%)	18 (17.8%)	101 (100%)
Smoker	6 (5.9%)	36 (35.6%)	59 (58.4%)	101 (100%)
PAF**	11 (10.9%)	72 (71.3%)	18 (17.8%)	101 (100%)
Valve Surgery	25 (24.8%)	58 (57.4%)	18 (17.8%)	101 (100%)

B

	≥ 60	< 60	Unknown	Total
Age	56 (55.4%)	27 (26.7%)	18 (17.8%)	101 (100%)

C

	Asian	Black	Hispanic /Latino	Indian	NAV	South American	White	Unknown	Total
Ethnicity	7 (6.9%)	1 (1.0%)	10 (9.9%)	1 (1.0%)	1 (1.0%)	1 (1.0%)	50 (49.5%)	30 (29.7%)	101 (100%)

* POAF: Post-Operative Atrial Fibrillation

* * PAF: Paroxysmal Atrial Fibrillation

Table 4-1. Demographic and clinical characteristic of 101 enroll patients that underwent CABG. **A.** Information of POAF, sex, PAF, valve surgery condition. 83 patients have data for post-operative atrial fibrillation, sex, paroxysmal atrial fibrillation, and valve surgery condition; the smoking history of 42 patients was known. POAF occurrence was at least 34.6%. **B.** Patient's age distribution. 83 patients have age data and a large portion of the patients who had the surgery were older than 60. **C.** Ethnicity data for enrolled patients. Data is present for 71 of the patients.

Table 4-2 Associated cardiovascular-related GO & KEGG pathway terms with involved CpG sites

Term	Gene/Promoter involved	Condition
Hypertrophic cardiomyopathy (HCM)	ACTB;CACNA1C;ITGB7;TGFB1; TPM3;ITGA11;PRKAG2	Hypervariable CpGs
Arrhythmogenic right ventricular cardiomyopathy (ARVC)	ACTB;ACTN4;CTNNA1;ITGA9;TCF7;CACNA2D2;LEF1	Hypervariable CpGs
Heart valve development	PRDM1;TGFB1	POAF
Transforming growth factor beta receptor signaling pathway involved in heart development	TGFB1	POAF
Hypertrophic cardiomyopathy (HCM)	ITGA9;ITGB7;TGFB1;PRKAG2	Valve surgery
Dilated cardiomyopathy	ITGA9;ITGB7;TGFB1	Valve surgery
Regulation of vascular permeability	ABR;AZU1;SRC;TGFB1	Valve surgery

Table 4-2. Associated cardiovascular-related GO & KEGG pathway terms with involved CpG sites. 7 terms related to the cardiovascular system for hypervariable CpG sites and significant hypervariable CpG sites for patient group separation (POAF and valve surgery condition) were shown. TCG-β1 was covered in most of the terms (shown in red). Also, genes in integrin family were appeared in more than half of the terms (show in blue).

Table 4-3 Clinical and demographic data that may affect the POAF occurrence

	Yes	No	Unknown	Total	P value
POAF	35 (34.6%)	48 (47.5%)	18 (17.8%)	101 (100%)	
Male	60 (59.4%)	23 (22.8%)	18 (17.8%)	101 (100%)	0.9279
PAF	11 (10.9%)	72 (71.3%)	18 (17.8%)	101 (100%)	0.0546
Valve Surgery	25 (24.8%)	58 (57.4%)	18 (17.8%)	101 (100%)	0.5625
Age	Range from 18 to 84, Mean=62.39				0.0191

Table 4-3. Clinical and demographic data that may affect the POAF occurrence. Logistic regression was applied to calculate the prediction power of sex, PAF, valve surgery condition and age. Based on the p-value cutoff, age may be used as the factor for the incident of POAF.

References

1.   Mathew JP, Fontes ML, Tudor IC, Ramsay J, Duke P, Mazer CD, Barash PG, Hsu PH, Mangano DT, Investigators of the Ischemia R, Education F, Multicenter Study of Perioperative Ischemia Research G. A multicenter risk index for atrial fibrillation after cardiac surgery. *Jama*. 2004;291:1720-1729

2.   Echahidi N, Pibarot P, O'Hara G, Mathieu P. Mechanisms, prevention, and treatment of atrial fibrillation after cardiac surgery. *Journal of the American College of Cardiology*. 2008;51:793-801

3.   Bird AP, Wolffe AP. Methylation-induced repression--belts, braces, and chromatin. *Cell*. 1999;99:451-454

4.   Jones PA. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nature reviews. Genetics*. 2012;13:484-492

5.   Cedar H, Bergman Y. Linking DNA methylation and histone modification: Patterns and paradigms. *Nature reviews. Genetics*. 2009;10:295-304

6.   Robertson KD. DNA methylation and human disease. *Nature reviews. Genetics*. 2005;6:597-610

7.   Gilsbach R, Preissl S, Gruning BA, Schnick T, Burger L, Benes V, Wurch A, Bonisch U, Gunther S, Backofen R, Fleischmann BK, Schubeler D, Hein L. Dynamic DNA methylation orchestrates cardiomyocyte development, maturation and disease. *Nature communications*. 2014;5:5288

8.   Chen H, Orozco LD, Wang J, Rau CD, Rubbi L, Ren S, Wang Y, Pellegrini M, Lusis AJ, Vondriska TM. DNA methylation indicates susceptibility to isoproterenol-induced cardiac pathology and is associated with chromatin states. *Circulation research*. 2016;118:786-797

9.   Meder B, Haas J, Sedaghat-Hamedani F, Kayvanpour E, Frese K, Lai A, Nietsch R, Scheiner C, Mester S, Bordalo DM, Amr A, Dietrich C, Pils D, Siede D, Hund H, Bauer A, Holzer DB, Ruhparwar A, Mueller-Hennessen M, Weichenhan D, Plass C, Weis T, Backs J, Wuerstle M, Keller A, Katus HA, Posch AE. Epigenome-wide association study identifies cardiac gene patterning and a novel class of biomarkers for heart failure. *Circulation*. 2017;136:1528-1544

10.   Chen PY, Ganguly A, Rubbi L, Orozco LD, Morselli M, Ashraf D, Jaroszewicz A, Feng S, Jacobsen SE, Nakano A, Devaskar SU, Pellegrini M. Intrauterine calorie restriction affects placental DNA methylation and gene expression. *Physiological genomics*. 2013;45:565-576

11.	Guo W, Fiziev P, Yan W, Cokus S, Sun X, Zhang MQ, Chen PY, Pellegrini M. Bs-seeker2: A versatile aligning pipeline for bisulfite sequencing data. *BMC genomics*. 2013;14:774

12.	Akalin A, Franke V, Vlahovicek K, Mason CE, Schubeler D. Genomation: A toolkit to summarize, annotate and visualize genomic intervals. *Bioinformatics*. 2015;31:1127-1129

13.	Assenov Y, Muller F, Lutsik P, Walter J, Lengauer T, Bock C. Comprehensive analysis of DNA methylation data with rnbeads. *Nature methods*. 2014;11:1138-1140

14.	Falcon S, Gentleman R. Using gostats to test gene lists for go term association. *Bioinformatics*. 2007;23:257-258

15.	Ioshikhes IP, Zhang MQ. Large-scale human promoter mapping using cpg islands. *Nature genetics*. 2000;26:61-63

16.	Deaton AM, Bird A. Cpg islands and the regulation of transcription. *Genes & development*. 2011;25:1010-1022

17.	Straussman R, Nejman D, Roberts D, Steinfeld I, Blum B, Benvenisty N, Simon I, Yakhini Z, Cedar H. Developmental programming of cpg island methylation profiles in the human genome. *Nature structural & molecular biology*. 2009;16:564-571

18.	Wagner JR, Busche S, Ge B, Kwan T, Pastinen T, Blanchette M. The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts. *Genome biology*. 2014;15:R37

19.	Letterio JJ, Roberts AB. Regulation of immune responses by tgf-beta. *Annual review of immunology*. 1998;16:137-161

20.	Dobaczewski M, Chen W, Frangogiannis NG. Transforming growth factor (tgf)-beta signaling in cardiac remodeling. *Journal of molecular and cellular cardiology*. 2011;51:600-606

21.	Burstein B, Nattel S. Atrial fibrosis: Mechanisms and clinical relevance in atrial fibrillation. *Journal of the American College of Cardiology*. 2008;51:802-809

22.	Seko Y, Nishimura H, Takahashi N, Ashida T, Nagai R. Serum levels of vascular endothelial growth factor and transforming growth factor-beta1 in patients with atrial fibrillation undergoing defibrillation therapy. *Japanese heart journal*. 2000;41:27-32

23.	Blobe GC, Schiemann WP, Lodish HF. Role of transforming growth factor beta in human disease. *The New England journal of medicine*. 2000;342:1350-1358

24. Verheule S, Sato T, Everett Tt, Engle SK, Otten D, Rubart-von der Lohe M, Nakajima HO, Nakajima H, Field LJ, Olgin JE. Increased vulnerability to atrial fibrillation in transgenic mice with selective atrial fibrosis caused by overexpression of tgf-beta1. *Circulation research*. 2004;94:1458-1465

25. Schultz Jel J, Witt SA, Glascock BJ, Nieman ML, Reiser PJ, Nix SL, Kimball TR, Doetschman T. Tgf-beta1 mediates the hypertrophic cardiomyocyte growth induced by angiotensin ii. *The Journal of clinical investigation*. 2002;109:787-796

26. Kostin S, Klein G, Szalay Z, Hein S, Bauer EP, Schaper J. Structural correlate of atrial fibrillation in human patients. *Cardiovascular research*. 2002;54:361-379

27. Lee KW, Everett THt, Rahmutula D, Guerra JM, Wilson E, Ding C, Olgin JE. Pirfenidone prevents the development of a vulnerable substrate for atrial fibrillation in a canine model of heart failure. *Circulation*. 2006;114:1703-1712

28. Giancotti FG, Ruoslahti E. Integrin signaling. *Science*. 1999;285:1028-1032

29. Parkin J, Cohen B. An overview of the immune system. *Lancet*. 2001;357:1777-1789

30. Ross RS, Pham C, Shai SY, Goldhaber JI, Fenczik C, Glembotski CC, Ginsberg MH, Loftus JC. Beta1 integrins participate in the hypertrophic response of rat ventricular myocytes. *Circulation research*. 1998;82:1160-1172

31. Chen C, Li R, Ross RS, Manso AM. Integrins and integrin-related proteins in cardiac fibrosis. *Journal of molecular and cellular cardiology*. 2016;93:162-174

32. Henderson NC, Arnold TD, Katamura Y, Giacomini MM, Rodriguez JD, McCarty JH, Pellicoro A, Raschperger E, Betsholtz C, Ruminski PG, Griggs DW, Prinsen MJ, Maher JJ, Iredale JP, Lacy-Hulbert A, Adams RH, Sheppard D. Targeting of alphav integrin identifies a core molecular pathway that regulates fibrosis in several organs. *Nature medicine*. 2013;19:1617-1624

33. Reed NI, Jo H, Chen C, Tsujino K, Arnold TD, DeGrado WF, Sheppard D. The alphavbeta1 integrin plays a critical in vivo role in tissue fibrosis. *Science translational medicine*. 2015;7:288ra279

34. Patsenker E, Popov Y, Stickel F, Jonczyk A, Goodman SL, Schuppan D. Inhibition of integrin alphavbeta6 on cholangiocytes blocks transforming growth factor-beta activation and retards biliary fibrosis progression. *Gastroenterology*. 2008;135:660-670

35. Turner CA, Jr., Mack DH, Davis MM. Blimp-1, a novel zinc finger-containing protein that can drive the maturation of b lymphocytes into immunoglobulin-secreting cells. *Cell*. 1994;77:297-306

36.   Kallies A, Hawkins ED, Belz GT, Metcalf D, Hommel M, Corcoran LM, Hodgkin PD, Nutt SL. Transcriptional repressor blimp-1 is essential for t cell homeostasis and self-tolerance. *Nature immunology*. 2006;7:466-474

37.   Robertson EJ, Charatsi I, Joyner CJ, Koonce CH, Morgan M, Islam A, Paterson C, Lejsek E, Arnold SJ, Kallies A, Nutt SL, Bikoff EK. Blimp1 regulates development of the posterior forelimb, caudal pharyngeal arches, heart and sensory vibrissae in mice. *Development*. 2007;134:4335-4345