UNIVERSITY OF CALIFORNIA

Los Angeles

Chromatin Architectural Dynamics in Cardiovascular Disease

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Molecular, Cellular, and Integrative Physiology

by

Douglas Chapski

2019

ABSTRACT OF THE DISSERTATION

Chromatin Architectural Dynamics in

Cardiovascular Disease

by

Douglas Chapski

Doctor of Philosophy in Molecular, Cellular, and Integrative Physiology

University of California, Los Angeles, 2019

Professor Thomas M. Vondriska, Chair

The chromatin architectural rearrangements that permit disease gene expression are just beginning to come to light. Distinct levels of chromatin organization are needed to maintain a healthy transcriptome, from the histone octamer that forms nucleosomes (the functional unit of chromatin) to chromosome territories that demarcate large swaths of the nucleus. An integrative picture of how each level of chromatin contributes towards healthy and disease gene expression has eluded us until chromosome conformation capture followed by high-throughput sequencing paved the way for deeper study of how chromatin features, such as significant chromosomal interactions, topologically associating domains, A/B compartmentalization, and enhancer-gene interactions all contribute towards gene regulation at a global scale. Heart failure is a syndrome

characterized, in part, by a dysregulated gene expression program. We hypothesized that chromatin structure becomes deranged during heart failure, and we found this to be the case at multiple levels of chromatin organization. In addition, we found that healthy cardiac myocyte chromatin structure permits its organ-specific gene regulation program. This dissertation 1) summarizes the cardiovascular epigenetics work in the field; 2) reports our findings from our chromosome conformation studies in hearts that underwent pressure overload and those that underwent knockout of the chromatin structural protein CTCF to understand how pathology influences chromatin structure; and 3) reports our investigation into the chromatin architectural organization of healthy cardiac myocytes when compared to healthy liver, in an effort to understand how normal epigenomes are organized in three dimensions. These studies open avenues for future mechanistic cardiovascular epigenomics studies, as well as for exploration of therapeutic treatment of chromatin to promote healthier gene expression programs.

The dissertation of Douglas Chapski is approved.

Matteo Pellegrini

James N. Weiss

Yibin Wang

Thomas M. Vondriska, Committee Chair

University of California, Los Angeles

2019

This thesis is dedicated to my grandmother,

Vitalina de Oliveira Carvalho,

who succumbed to heart failure during my PhD,

and to the patients who make

our studies worth the journey.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First, I would like to thank my PhD advisor, Dr. Thomas M. Vondriska, for his support during my doctoral studies. He took a chance on me when I was first learning to analyze high-dimensional data, and it paid off: our work has become something to be appreciated by those who wish to understand epigenomic mechanisms of heart failure. Tom has been a key figure in the development of my thinking style, and his guidance with study design and experimental methodologies has proven invaluable. Importantly, he has taught me to navigate the world with poise and courtesy.

Here I acknowledge the profound impact that the members Vondriska Lab have made on me over the years. In particular, Dr. Manuel Rosa-Garrido has been a mentor and friend who never doubted my technical abilities, even when conceptual and computational challenges appeared insurmountable. His relentless pursuit of chromatin structural insights is to be admired. Maximilian Cabaj has provided a new perspective on technical procedures in the lab, and has imparted wholesome cultural wisdom along the way. Dr. Emma Monte taught me that science is not just about the results, but also about the questions that are asked. She can find holes in any ostensibly airtight argument, and her sustained enthusiasm for science continues to be an inspiration. Dr. Elaheh Karbassi has been a beacon of knowledge with respect to chromatin structure in the cardiac myocyte, and has shown that one must have grit and a healthy supply of tenacity to make admirable achievements. Todd H. Kimball has provided consistent technical support with wet lab experiments, and offered me the unique and exciting opportunity to help him with a

computational investigation of how bird genomes are regulated during his initial graduate work. His kindness and commitment to science are refreshing. All other lab members, past and present, have made true marks on me as a scientist and as a citizen of the world.

Here I thank the journals *Circulation Research*, *Circulation*, and *Frontiers in Cardiovascular Medicine*, who allow us to use our open access, published information in this dissertation. The *Circulation Research* paper (Chapter 1, published in *Circulation Research*. 2018 May 25;122(11):1586-1607, PMID: 29798902) involved Manuel Rosa-Garrido and Thomas M. Vondriska, who worked on concepts, writing and figure making. The *Circulation* paper (Chapter 2, published in *Circulation*. 2017 Oct 24;136(17):1613-1625, PMID: 28802249) involved Manuel Rosa-Garrido, Anthony D. Schmitt, Todd H. Kimball, Elaheh Karbassi, Emma Monte, Enrique Balderas, Matteo Pellegrini, Tsai-Ting Shih, Elizabeth Soehalim, David Liem, Peipei Ping, Niels J. Galjart, Shuxun Ren, Yibin Wang, Bing Ren, and Thomas M. Vondriska, who helped with concepts, wet lab experiments, bioinformatics, and writing. The *Frontiers* paper (Chapter 3, published in *Frontiers in Cardiovascular Medicine*. 2019 Jan 15;5:186, PMID: 30697540) involved Manuel Rosa-Garrido, Nan Hua, Frank Alber and Thomas M. Vondriska, who helped with concepts, figures, and writing.

I am also grateful for the Eugene V. Cota-Robles Fellowship from the University of California, as well as the Ruth L. Kirschstein National Research Service Award

T32HL69766. The Vondriska Lab is funded by the National Institutes of Health and the David Geffen School of Medicine at UCLA.

My PhD Committee is composed of four seasoned scientists. In addition to my PhD mentor, Dr. Matteo Pellegrini provided intellectual and technical support with genomics experiments and data analyses on short notice. Dr. James N. Weiss showed me how a true passion for science, a sincere compassion for patients, and a deep sense of creativity can coalesce to reveal profound insights within the field of cardiovascular systems biology. Dr. Yibin Wang provided unending collaborative support for my scientific endeavors and showed me that innovative questions lead to remarkable findings in biomedical research.

I would also like to thank my family, friends, and former colleagues for inspiring me to contribute towards the advancement of science, which is larger than the sum of its parts. My parents, while not scientists, understand the importance of working towards a goal that makes a difference. I will always cherish their patience and supportive voices during each phone call. Dr. Christoph D. Rau has shown me how elegant statistical algorithms can illuminate biological datasets. Dr. William D. Barshop has been a true friend and colleague over the years, and his statistical discussions and exemplary passion for science still resonate with me. Dr. Rafael S. Demarco has provided consistent scientific and life advice, and his commitment to basic research is evident in all that he does. Drs. Ray Chui and Pradeep Rajendran have taught me new ways to think about heart disease and consistently remind me that the heart can throw some real curveballs. Dr. Michael S.

Joseph has been an inspiration throughout his PhD work and during his early years as a professor.

Most importantly, I express my deepest gratitude to my wife Brenna, and my daughter Eloise, for showing me that there exists true love in this world, and for patiently tolerating the scientific process. They are my lighthouse.

BIOGRAPHICAL SKETCH

PhD in Molecular, Cellular & Integrative Physiology | Expected Spring 2019
University of California, Los Angeles

Bachelor of Science in Cell & Molecular Biology and Spanish | 2012, *cum laude*
Tulane University

POSITIONS AND EMPLOYMENT

2008-2009 CURE Student, Dept. of Structural Biology, Dana-Farber, Boston, MA
2008-2009 Volunteer Student, Dept. of Biochemistry, LSUHSC, New Orleans, LA
2009-2012 Undergrad Laboratory Technician, Tulane University, New Orleans, LA
2010-2011 SHURP Student, Dept. of Cell Biology, Harvard Medical, Boston, MA
2012-2012 UCLA Competitive Edge Scholar, UCLA, Los Angeles, CA
2014-2015 Teaching Assistant, Dept. of Life Sciences, UCLA, Los Angeles, CA
2013-2019 Graduate Student, Dept. of Anesthesiology, UCLA, Los Angeles, CA

HONORS

2008-2012 Dresser Memorial Scholarship, Tulane University (~$30k/year)
2009 Phi Eta Sigma National Honor Society
2012 Louisiana Alliance for Minority Participation Scholarship, Tulane University
2012 Nelson Betts Roltsch Memorial Scholarship, Tulane University
2012 Cell and Molecular Biology Prize, Tulane University
2012 UCLA Competitive Edge Scholar (Funded by NSF)
2012 Eugene V. Cota-Robles Fellowship, Round 1 (Year 1 of PhD)
2013 NSF Graduate Research Fellowship Honorable Mention
2015 UCLA Molecular, Cellular, & Integrative Physiology Retreat Poster Winner
2016 UCLA Doctoral Student Travel Grant
2016 Eugene V. Cota-Robles Fellowship, Round 2 (Year 5 of PhD)
2017-2018 UCLA Cardiovascular Symposium Poster Winner (both years)

CONTRIBUTION TO SCIENCE

1. **Signal transduction events in disease.** During undergraduate work, I received intensive training in signaling biology, where I worked on signaling and metabolism experiments in the mTOR pathway. Strikingly, we found that the mTORC1 pathway regulates glutamine metabolism via modulation of a sirtuin called SIRT4. This investigation resulted in a publication in the journal *Cell*, and an opportunity to write an editorial about the mTOR pathway in the journal *Circulation Research* during my PhD studies.
    a. Csibi A, Fendt S-M, Li C, Poulogiannis G, Choo AY, **Chapski DJ**, Jeong SM, Dempsey JM, Parkhitko A, Morrison T, Henske EP, Haigis MC, Cantley LC, Stephanopoulos G, Yu J, Blenis J. The mTORC1 Pathway Stimulates Glutamine Metabolism and Cell Proliferation by Repressing SIRT4. Cell. Elsevier Inc.; 2013 May 9;153(4):840–54.
    b. **Chapski DJ**, Monte E, Vondriska TM. Positive feedback in cardioprotection: can more mechanism lead to translation? Circ Res. 2014 Apr 11;114(8):1225–7.

2. **Chromatin architecture in cardiovascular disease.** During my PhD studies, I have been investigating the role of chromatin structure on gene expression. Our goal is to understand the mechanisms that underpin the global shift in chromatin structure during disease, and to determine whether chromatin structure globally influences gene expression after cardiac perturbation. Our experiments, in addition to other chromatin investigations by the lab, revealed that transcriptional changes in disease are accompanied by chromatin architectural rearrangements, and that perturbation of myocardium via pressure overload causes aberrant chromatin conformational shifts. We also published a deep study into the interactions that contribute to organ-specific nuclear architectures.
    a. 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 Oct 24;136(17):1613-1625.
    b. Rosa-Garrido M, **Chapski DJ**, Vondriska TM. Epigenomes in Cardiovascular Disease. Circ Res. 2018 May 25;122(11);1586–1607.
    c. **Chapski DJ**, Rosa-Garrido M, Hua N, Alber F, Vondriska TM. Spatial Principles of Chromatin Architecture Associated with Organ-Specific Gene Regulation. Frontiers in Cardiovascular Medicine. 2019 Jan 15;5:186 .

3. **Aberrant gene expression in cardiovascular disease.** A hallmark of heart failure is a pattern of dysregulated gene expression termed the “fetal gene program.” After inspecting the Hybrid Mouse Diversity Panel, a panel of inbred and genetically distinct mouse strains, we observed that canonical “fetal genes” do not behave consistently across genetic backgrounds. This has major translational implications. Another study we have worked on, in collaboration with the lab of Dr. Yibin Wang, is the contribution of RBFox1 to regulation of cardiac hypertrophy and failure.
    a. Karbassi E, Monte E, **Chapski DJ**, Lopez R, Rosa Garrido M, Kim J,Wisniewski N, Rau CD, Wang JJ, Weiss JN, Wang Y, Lusis AJ, Vondriska TM. Relationship of disease-associated gene expression to cardiac phenotype is buffered by genetic diversity and chromatin regulation. Physiol Genomics. 2016 Aug 1;48(8):601–15.
    b. Gao C, Ren S, Lee J, Qiu J, **Chapski DJ**, Rau CD, Zhou Y, Abdellatif M, Nakano A, Vondriska TM, Xiao X, Fu XD, Chen JN, Wang Y. RBFox1-mediated RNA splicing regulates cardiac hypertrophy and heart failure. J Clin Invest. 2016 Jan;126(1):195–206.

Completed Research Support

Eugene V. Cota-Robles Fellowship 09/2012 – 07/2013 and 09/2016 – 07/2017
Provided full tuition, stipend, and fees for the 1st and 5th year of PhD studies at UCLA

T32HL69766, Iruela-Arispe (PI) 07/01/2013 – 07/01/2016
Ruth L. Kirschstein National Research Service Award

Preface

Heart failure occurs when the organ cannot supply enough oxygen and nutrients to the body via systemic circulation. The disease affected ~6.2 million people over 20 years of age in the United States between 2013 and 2016 and its monetary cost is estimated to increase to $69.8 billion by 2030 (from $30.7 billion in 2010)[1]. During cardiac stress by high blood pressure, catecholamines, and other stimuli, left ventricular hypertrophy is initially observed, followed by maladaptive remodeling that causes dilation of the chamber and thinning of the ventricular wall, leading to subsequent reduction in ejection fraction[2]. Concomitantly, cardiomyocytes (the post-mitotic functional unit of the organ) undergo dysregulated gene expression, termed the "fetal gene program"[3], which is reminiscent of the gene expression program observed during cardiomyocyte development. Unknown, however, is the comprehensive genomic picture of *how* these genes become dysfunctional from the normal, quiescent state on a global scale. In addition to locus-specific gene regulation by transcriptional activators or repressors, it is understood that genes spatially move into and out of chromatin territories that permit transcriptional activity[4]. Because structure contributes strongly to function within many biological contexts, we seek to understand the global structural rearrangements that occur with heart failure, to better explain the dysregulation at the transcriptomic scale, and to complement other chromatin structural analyses that have been done in the heart.

Many groups have studied gene regulation at the locus level, and my aim is to study how higher-order chromatin structure is involved in cardiac pathology. My projects investigate chromatin structure at distinct feature levels to understand how genes become dysfunctional during heart failure. All of these studies incorporate chromosome

conformation capture followed by high-throughput sequencing[5], also known as Hi-C, to understand chromatin structural dynamics during disease, and some examples of investigated structural features are given below: Topologically associating domains (TADs) are chromatin features that tend to interact with each other more than with other regions, and whose boundaries are occupied by the chromatin structural protein CTCF and prevent the spread of heterochromatin[6]. Our studies in Chapter 2 reveal that, across the genome, TAD boundaries change in strength with cardiac perturbation (either pressure overload or cardiac-specific CTCF-knockout). The results from the study help to explain why failing cardiomyocytes have such drastic differences at the transcriptome level when compared to healthy cardiac myocytes. A/B compartmentalization is another level of chromatin structure that changes with disease (compartment A tends to contain more active regions while compartment B tends to contain more inactive regions, although this is a trend and not a rule). Chapter 2 shows how pressure overload induced heart failure and cardiomyocyte-specific CTCF-knockout cause disruption in compartmentalization (i.e., changes from compartment A to B and vice versa), and subsequent differential gene expression within those changing compartments. Chapter 3 also investigates A/B compartment dynamics by comparing the cardiomyocyte genome to the liver genome. These changes are remarkable and can help to explain why organs have their own distinct gene expression phenotypes despite having the same genomic sequence as all other cells in the body. Enhancers are another chromatin feature thought to regulate gene expression by recruiting genomic machinery to, and thereby activating, gene promoters in three-dimensional space[7]. Canonically, studies have been done on the closest genes to putative enhancers, which is not necessarily what is observed in reality:

our work in Chapter 2 uses Hi-C to show that enhancers interact with many genes in space, and they may not behave as expected on a global scale. This provides a rationale for studying gene regulation with chromosome conformation capture based techniques to better understand how genes become dysregulated with disease.

This dissertation begins with a published review on epigenomics in the heart (Chapter 1). Rigorous science builds upon previously published studies; many of those most relevant to our work are referenced in Chapter 1 (Table 1-1). By recognizing and integrating the concepts from prior work in the field of cardiovascular epigenetics, we are better prepared to make our own scientific endeavors. We seek to get the big picture and then highlight specific features, integrating our novel insights with the large body of work done by others. The next chapter describes our study of chromatin structural dynamics in a pressure overload model of heart failure and in a CTCF-knockout model of chromatin disruption (Chapter 2). The final chapter is a published investigation into the differences in chromatin structure between cardiomyocytes and liver cells, to better explain how genomes can be organized for organ-specific function (Chapter 3). Lastly, a Future Directions section explains how, with the immense body of chromatin work done in the heart, we can begin to integrate vast bodies of data and push the field forward with innovative technologies and insightful questions to better understand genomic dysfunction during heart failure.

Preface: Bibliography

1. Benjamin, E. J. *et al.* Heart Disease and Stroke Statistics—2019 Update: A Report From the American Heart Association. *Circulation* **139,** e146–e603 (2019).
2. Kehat, I. & Molkentin, J. D. Molecular pathways underlying cardiac remodeling during pathophysiological stimulation. *Circulation* **122,** 2727–35 (2010).
3. 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,** 331–43 (2007).
4. Rosa-Garrido, M., Chapski, D. J. & Vondriska, T. M. Epigenomes in Cardiovascular Disease. *Circ. Res.* **122,** 1586–1607 (2018).
5. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326,** 289–93 (2009).
6. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485,** 376–380 (2012).
7. Pennacchio, L. A., Bickmore, W., Dean, A., Nobrega, M. A. & Bejerano, G. Enhancers: five essential questions. *Nat. Rev. Genet.* **14,** 288–95 (2013).

Chapter 1: Epigenomes in Cardiovascular Disease

Manuel Rosa Garrido, Douglas J. Chapski, Thomas M. Vondriska

[This research was originally published by Rosa-Garrido *et al.* Epigenomes in Cardiovascular Disease. *Circulation Research*. 2018 May 25;122(11):1586-1607. PMID: 29798902 © The Authors.]

Abstract

If unifying principles could be revealed for how the same genome encodes different eukaryotic cells *and* for how genetic variability and environmental input are integrated to impact cardiovascular health, grand challenges in basic cell biology and translational medicine may succumb to experimental dissection. A rich body of work in model systems has implicated chromatin modifying enzymes, DNA methylation, noncoding RNAs and other transcriptome-shaping factors in adult health and in the development, progression and mitigation of cardiovascular disease. Meanwhile, deployment of epigenomic tools, powered by next generation sequencing technologies in cardiovascular models and human populations, has enabled description of epigenomic landscapes underpinning cellular function in the cardiovascular system. This essay aims to unpack the conceptual framework in which epigenomes are studied and to stimulate discussion on how principles of chromatin function may inform investigations of cardiovascular disease and the development of new therapies.

Key words: cardiac; vascular; epigenetics; genetics; chromatin

The tracks of a train point where it is going—they also reveal where it has been. The tracks of the leopard in the snow do only the latter. Which is the case for epigenetic processes in cardiovascular health and disease? Chromatin enables bespoke storage and retrieval of genetic information across the hundreds of cells in a multicellular organism. Chromatin is also perhaps the largest, in terms of physical size and list of component parts, and most functionally diverse molecular structure in existence, enabling the vast diversity of the multicellular world. The goals of this essay are: (1) to review recent advancements in the concepts of epigenomics; (2) to summarize the evidence that epigenomic processes participate in normal and diseased cardiovascular physiology; and (3) to stimulate discussion about the prospects for novel basic science and translational investigations of epigenomes in the cardiovascular system.

In development, the setting in which epigenesis has been most extensively examined, the task of epigenetic processes is to guide the unidirectional differentiation of cells. In adulthood, epigenetic mechanisms provide stability, maintaining the blueprints laid down in development and resisting environmental stresses and stochastic intracellular changes. In cardiovascular disease, an area in which epigenetic processes are increasingly appreciated to operate and in which insights into human health are the most actionable, the task, teleologically, of chromatin is not all that clear. Are the epigenetic processes at work in the cell during disease always combating the insult, attempting to restore and preserve healthy adult phenotypes? Have epigenetic processes been hijacked by evolution to contribute to cell survival decisions? What are the operational principles of epigenetics and how are they distinguished from gene

regulation? For this essay, we define an epigenome as a genome plus everything binding to and modifying it (please see Supplemental Glossary).

Evidence has recently accumulated that epigenomic processes play a central role in cardiovascular disease, including: (1) genetic and pharmacologic gain and loss of function studies in the cardiovascular system targeting individual histone modifying enzymes, chromatin remodelers, chromatin structural components and regulatory RNAs can induce or prevent pathology; (2) descriptive epigenomic investigations from humans and animal models demonstrate that widespread reprogramming of histone modifications, DNA methylation and protein binding occurs during the development of disease; (3) data has arisen that indicates epigenomic processes are interconnected with other cellular events (such as metabolism, differentiation and cell death) that are deranged in disease and that modulation of these epigenetic processes may hold potential for therapeutic intervention. Let us now examine recent advancements in our understanding of how chromatin controls gene expression and consider the implications for cardiovascular disease.

Structural components of chromatin and their regulation in the cardiovascular system

The epigenome has two fundamental components: the structural features of chromatin and the enzymes, RNAs, small molecules and processes that modify these features. For example, a nucleosome is the chromatin structural unit, but an ATP-dependent chromatin remodeler is an enzyme complex that modifies this structural unit. Here is how you make a nucleosome[1]: bind histone H3 to histone H4 and combine two

of these dimers to form a tetramer; then combine this tetramer with histone H2A and histone H2B, twice, and embrace this octamer in turn with ~145-147 bp of DNA. Crystal structures[2] of the nucleosome in 1997 revealed this octameric protein complex to be assembled through histone fold dimers of H2A with H2B and H3 with H4. The intervening years have seen the discovery of multiple histone isoform variants that can exert structural changes at the atomic level on the functional properties of the nucleosome: one example is histone H2A.z, an H2A variant associated with active transcription, which operates by altering the interfaces between H2A-H2B dimers, as well as amongst the entire tetramer, partially destabilizing the nucleosome core (thereby facilitating nucleosome eviction during transcription).[3] Another example is centromeric histone H3, called CENP-A, which replaces H3 in centromeric nucleosomes, altering the interactions amongst histones but maintaining the octameric structure.[4] Mammalian genomes encode multiple versions of core histones, with each family hosting multiple variants of unknown significance. These histone isoform variants have been shown to participate in development and disease through altered roles in DNA replication, chromosome packaging and DNA accessibility.[5, 6] Proteomics experiments have demonstrated variable expression of these isoforms in different cell types, including in the heart,[7] and the stoichiometry of histones changes with disease, such as in pressure overload hypertrophy.[8] In general, however, the nucleosome is amongst the most conserved protein structures known.

Where nucleosomes bind along the genome is influenced—certainly in reconstituted systems but likely also *in vivo* to some degree—by primary DNA sequence, with poly dA:dT rich regions being comparably depleted of nucleosomes, due in part to

the biophysical properties of these regions that resist bending necessary for nucleosome binding.[9] An additional class of chromatin structural element includes non-nucleosome chromatin structural proteins, such as histone H1 family proteins, CTCF, and high mobility group (HMG) proteins. Histone H1, the so-called linker histone (also known as histone H5), is not part of the nucleosome core particle, but may participate in the formation of higher order chromatin fibers through interactions with nucleosome histones and DNA.[10] Structures of nucleosomes with H1 (which has been termed the chromatosome[11]) show the latter nestled just outside of the core,[12, 13] making independent contacts with DNA, but whether linker histones are obligatory components of a structural feature in eukaryotic chromatin remains a matter of debate. Observations from a variety of eukaryotic systems implicate linker histone H1 in genomic functions including DNA repair/stability, replication, and transcription.[14] CTCF and HMG proteins each have likewise been attributed a wide range of possible functions, based mostly on gain/loss of function studies in cell systems. As a general mechanism, CTCF binds DNA and facilitates chromatin looping (i.e. the formation of semi-stable long range intra-chromosomal interactions, spanning kilobases);[15] specificity may arise in different cell types to facilitate regulatory interactions. HMG proteins, of which there are many families, bind and stabilize distinct structural features in DNA, thereby contributing to high order chromatin anatomy and perhaps assemblies of nucleosomes. Interestingly, the HMGA family of these proteins has been implicated in cardiac disease, with both hetero- and homozygous knockout mice developing hypertrophy.[16]

How nucleosomes position along chromatin is largely specified, in multicellular eukaryotes, by ATP-dependent chromatin remodeling enzymes and the trio of protein

classes responsible for the so-called the histone code, those being writers, erasers and readers, that add, remove or interpret histone modifications, respectively. Although all these processes exist in other organelles and compartments (indeed many purported histone modifiers in fact interface with non-histone, non-nuclear proteins), the allegorical "grammar" terminology adopted for chromatin PTMs has been particularly helpful in conceptualizing experiments into how gene expression is regulated and particularly effective in biasing the interpretation of results.

Since their discovery and association with transcription or its inhibition, histone modifications have been the source of innumerable question-begging experiments and seemingly endless rounds of debate regarding their relationship to gene regulation.[17-19] Histone post-translational modifications were originally described in the 1960's[20] but only with the advent of chromatin immunoprecipitation and DNA sequencing (ChIP-seq; Figures 1-1 and 1-2) studies could the occupancy of these modifications be correlated, in a genome-wide manner in a single experiment, with gene transcription. Histone H3 has been particularly well characterized and shown to exhibit conserved regulatory behavior across species and cell type, with consensus existing about the roles of H3K27me3 in reversible gene silencing, H3K27ac in gene activation, H3K4me3 in gene activation, H3K9me3 in more lasting gene silencing and heterochromatin, although other histone isoforms, such as residue K20me on histone H4 (associated with gene silencing), exhibit evolutionarily conserved relationships to gene regulation. Histone modifications are associated with a variety of processes not related directly to transcription including replication, DNA repair, mitosis, and cell division,[5, 6, 21] although these processes tend to

be better understood in model organisms like yeast and not extensively studied in mammalian systems *in vivo*.

The histone code hypothesis (discussed in greater detail below), as originally articulated by Strahl and Allis,[17] states that a prescribed transcriptional or other genomic regulatory response arises, or is evoked, from a certain combination of histone PTMs *in vivo*. This hypothesis has tacitly established a framework in which to examine large amounts of ChIP-seq data, inspecting for regions of correlation between histone marks and gene expression. The resolution of most ChIP-seq "peaks" (~200-1,000 bp) is highly dependent on the informatics tools used to map reads and subsequently call peaks[22] (because protein occupancy on the genome makes a much smaller footprint, it is consequently often unknown exactly where a protein binds; the less commonly applied ChIP-exo[23] and X-ChIP-seq,[24] achieving ~25-50bp resolution, are exceptions). Thus, one does not know whether histone PTMs occupy the same nucleosome, which would be required for a histone PTM "reader" protein to distinguish a full, accurate code from a partial and/or inaccurate code (example of partial exception: chromatin digestion to single nucleosomes coupled with mass spectrometry proved coexistence of modifications on the same particle [in an antibody independent manner], which incidentally revealed that bivalency can occur at the single nucleosome level [i.e. different histone H3 copies in the same complex have opposing—that is one has an active and the other a repressive—PTMs], so-called asymmetrically modified nucleosomes[25]).

Histone writers and erasers are signal transduction processes that sometimes use histones as substrates. The concept of the reader, however, is an indispensable component of epigenetic language and one without obvious counterpart outside the

context of the chromatin (scaffolding proteins in signaling networks are a similar but distinct concept). To work as advertised—that is, to distinguish target region alpha from non-target beta, to bind alpha, and then to *do something*—readers must be able to simultaneously recognize and discriminate between histone PTMs, which has been shown in some model systems,[26] but has not been demonstrated to operate in mammals in a tissue-specific manner. Progress has recently been made in the area of heart failure, however, with the example of the BET bromodomain protein BRD4, whose inhibition with the small molecule JQ1 prevents hypertrophy and associated transcriptional changes[27, 28] and whose pharmacological inhibition reverses some of the fibrotic and deleterious structural remodeling in the wake of infarction or pressure overload (having no effect on physiological hypertrophy).[29] These studies have raised the intriguing possibility that epigenetic readers may be novel targets of epigenetic therapy in cardiovascular disease,[30] given their cell type specific expression patterns, ability to integrate the actions of multiple chromatin modifiers and their role to program gene expression with high fidelity at individual loci.

All four core histones are extensively post-translationally modified, with different histones (and different modifications) studied to varying extents and degrees of rigor and functional depth. The bane of accurate and reproducible proteomic mass spectrometry is that a truly unbiased experiment to detect PTMs will identify many, with no certainty about which ones are important versus otiose signaling noise. The most common answer to the question of which PTMs are important for chromatin regulation is some variation on *only the ones shown to be functionally involved in a phenotype.* Often the PTMs deemed most tantalizing are those previously published on in lower organisms and/or those for which

commercially available antibodies exist. In the chromatin world, these PTMs and the associated antibodies are the ones getting the most game time in ChIP-seq experiments. A quantitative example: the silencing mark histone H3 lysine 27 trimethylation, which enriches around gene regulatory regions, is represented at the time of this writing by 403 ChIP-seq datasets in ENCODE and ~2000-3000 publications in PubMed (an exact count of publications is tough given cavalier nomenclature); those numbers for H3 K27 acetylation, a euchromatic mark (note: on the same residue of the same protein), identified in 384 and ~500, respectively. In contrast, a recently identified succinylation on human histone H3 K79, also found to enrich in the transcription start sites of active genes, has one publication and one ChIP-seq dataset.[31] When examining a new PTM, no good antibody, and/or no tractable cell system for transgenesis of a tagged surrogate, means no genome-wide address book, which means, for the present, the given modification is assumed to be noise. That histone modifications cannot be unequivocally examined by genetic approaches (i.e. by mutating a residue in the gene and studying the protein) is the reason why they are alternatively so frustrating and interesting.

What about other changes in protein levels in the nucleus? The first attempts to understand proteomic remodeling of the cardiomyopathic nucleus were conducted in hamsters.[32] These early studies did not directly measure histone PTMs—although they (the PTMs) likely were reflected in the electrophoretic patterns reported in these investigations—because of the absence of accurate and quantitative mass spectrometry of proteins and PTMs. More recently, quantitative analysis of nuclear proteins[33] or chromatin-associated proteins in the heart revealed the proteins regulating cardiac epigenomes[34] and described changes in histone protein isoform stoichiometry[8] in the

setting of pressure overload hypertrophy, yet these studies did not characterize histone PTMs. Renal nuclear proteomes have also been explored,[35] with physiological implications for cardiovascular disease. Across eukaryotes, ~500 post-translational individual modifications have been identified just on the core nucleosome particle.[21] Very few of these have been identified, much less quantified and proficiently interrogated, in the cardiovascular system (most studies, to our knowledge, of histone PTM in the cardiovascular system rely on commercially available antibodies). Recent studies have employed large libraries of histone modifications, in some cases representing >100 different nucleosome combinations (note: these are discrete, tailored nucleosomes with known histone post-translational modifications, in contrast to the total list of histone modifications, for which it is rarely known whether the modifications are happening on the same nucleosome, let alone in the same copy of the histone, with the exception of proteomic studies on intact proteins[36]), to investigate the ability of particular histone PTM combinations to influence the activity of human ATP-dependent chromatin remodelers.[37] These libraries[38] allow for the effects of individual recombinant nucleosomes with designer histone modifications (and combinations of modifications) to be tested *in vitro* for their ability to influence a host of chromatin properties, such as the binding preferences for transcription factors or chromatin modifying enzymes.

Role of histone modifying enzymes in cardiovascular physiology and pathophysiology

Investigations into the enzymes responsible for depositing and removing histone modifications have demonstrated striking phenotypes in a range of cardiovascular

syndromes. Histone deacetylases (HDACs) are one of the most extensively studied families of histone-modifying enzymes in the cardiovascular system. HDACs consist of four families, each with distinct isoforms which in turn have distinct histone and—in some cases—non-histone targets, distinct cellular locations and distinct biological functions. Inhibition of HDACs has been shown pharmacologically (e.g. with trichostatin A or valproic acid) to prevent proliferation of vascular smooth muscle cells[39, 40] (with implications for atherosclerosis[41]), to attenuate hypertension,[42] to ameliorate ischemic/reperfusion injury and post-ischemic remodeling[43-46] and to block cardiac hypertrophy in the setting of heart failure.[47-49] Molecular dissection of these phenomena, particularly in the setting of cardiac growth have revealed HDACs to be powerful hypertrophic modulators: loss of HDAC9 leads to prodigious cardiac growth,[50] HDAC4 and 5 regulate CamKII-dependent gene regulation,[51, 52] HDAC2 regulates GSK3beta-Akt-dependent fetal gene activation in hypertrophy,[53] to name just a few examples. Indeed, the literature on the role of HDAC targeting by drugs or genetic manipulation in the cardiovascular system is sufficiently vast to devour entire, authoritative reviews articles.[41, 47, 54, 55]

Class III HDACs are also known as Sirtuins (Sirt) and have been shown to exert powerful effects in the cardiovascular system following their initial identification and association with longevity in yeast. There are seven Sirtuin isoforms, each with varied subcellular localization and some of which, such as Sirt1, 2, 6 and 7, localize to nucleus and thus are poised to use DNA-bound histones as substrates, including in the heart.[56] Multiple Sirt isoforms have been shown to facilitate DNA repair through homologous recombination. Unlike other HDACs, Sirtuins require nicotinamide adenine dinucleotide

(NAD) as a cofactor for activity. Intriguingly, these proteins have been shown to participate in an isoform specific manner in myocardial ischemia, oxidative stress, apoptotic cell death and cardiac hypertrophy.[56] Sirt isoforms have also been shown to exert salubrious effects on fatty acid, amino acid and glucose metabolism in mouse studies, which have led to clinical trials.[57] A search at clinicaltrials.gov with the keywords "cardiovascular disease" and "epigenetics" revealed 29 studies in the active or recently completed phase. These studies range in their employment of epigenetic measurements or interventions. For example, some describe measuring epigenetic endpoints in response to treatments targeting known cardiovascular risk factors in common CVD (e.g. "Epigenetic Reprogramming of Monocytes in Patients With Coronary Atherosclerosis", which measures epigenetic marks in the promoters of pro-inflammatory cytokines and chemokines in monocytes). Another example is a trial targeting rare genetic diseases with epigenetic manifestations in the cardiovascular system "X-chromosome Inactivation, Epigenetics and the Transcriptome", which measured DNA methylation, histone PTMs and coding/noncoding RNAs expression in blood, white cells and other tissues. Other examples investigate use of hypomethylating agents including 5'azacytidine and 5-aza-2'-deoxycytidine in adult patients with acute myeloid leukemia and atrial fibrillation ("Action of the Vidaza on the Atrial Fibrillation") or study DNA methylation as a marker for high blood pressure in the setting of pregnancy ("Observational Study of Epigenomic Dysregulation in Preeclampsia-Associated Chronic Hypertension"). The Human Epigenetic Drug Database (hedds.org) is a useful, searchable resource for information on epigenetic drugs, their targets and associated datasets including, when relevant, links to clinical trials. Several epigenetic drugs are FDA approved for clinical use, including

HDAC inhibitors and DNMT inhibitors [58] and several of these have been explored in cardiovascular disease (this review[59] also includes an extensive summary of miRNAs targeted for therapeutic manipulation by pharmaceutical companies).

Histone acetyltransferases (HATs) also constitute a large family of genes and have numerous nuclear and non-nuclear substrates (with type A being nuclear and type B being non-nuclear, mainly cytoplasmic) and have been characterized with varying degrees of specificity in cardiac and vascular cells.[55, 60, 61] One of the best characterized HATs is p300 which, in addition to its common residence at enhancer elements, has been shown to regulate genes that inhibit endothelial cell inflammation in the setting of atherosclerosis[62] and to attenuate salt-induced hypertensive heart failure[63] and agonist-induced cardiac hypertrophy[64] (the males absent on the first [MOF] HAT has similar anti-hypertrophic actions when overexpressed in the mouse[65]).

The histone methyltransferase (HMT) SET domain containing 2 (SETD2) was recently shown to be essential for myoblast differentiation in a process involving modification of histone H3K36 trimethylation.[66] SET and MYND containing HMT (Smyd; a family with 5 isoforms of varying tissue expressivity) family member Smyd1, which is restricted to striated muscle, has been shown to participate in cardiac phenotype through loss of function studies in the adult heart, which resulted in hypertrophy, dilation and de-repression of some cardiac disease genes.[67] Although Smyd1 localizes to the nucleus and interacts with chromatin in the adult heart[68] (and has been shown to regulate H3K4me3, an activating mark, in reconstituted systems;[69] mice with Smyd1 depletion in adulthood exhibited sustained H3K4me3 levels,[68] however, supporting the existence of alternative substrates for Smyd1 and indicating the existence of alternative proteins

capable of maintaining H3K4me3 in cardiac myocytes), a substantial portion of the protein is non-nuclear. This non-nuclear population of the protein has also been implicated in adult cardiac function (along with another family member Smyd2[70]) and heart development.[71-73] A different SET family protein, G9a (also known as euchromatic histone lysine methyltransferase 2 − EHMT2), has been shown by loss of function studies to play a key role in adult cardiac phenotype: inducible MerCreMer-dependent depletion resulted in cardiac hypertrophy, modestly depressed ejection fraction and fibrosis through a mechanism that involves targeted derangement of multiple histone methylation marks (including H3K9me2/3 and H3K27me3) at genes involved in cardiac function.[74] SET7 was shown in a human microvascular endothelial cell line to regulate target gene (IL-8 and HMOX1) expression in a H3K4me1-dependent and -independent manner.[75] Intriguingly, this process is tightly linked to glucose levels, demonstrating a metabolic sensing mechanism on chromatin.[76] On the removal side of methylation, genetic loss or augmentation of the histone demethylase JMJD2A respectively blocked or exacerbated cardiac hypertrophy in mice.[77]A recent pharmacological study showed that inhibition of histone methylation at H3K9 with the compound chaetocin (which targets the enzyme [SU(VAR)3-9] responsible for conversion of H3K9me2 to H3K9me3) attenuates some aspects of salt-induced cardiac dysfunction (survival rate, fractional shortening and fibrosis)—while not affecting pathologic gene regulation and only modestly impacting hypertrophic growth—in part through diminished H3K9me3 at repetitive elements and mitochondrial genes.[78]

The polycomb repressive complex (PRC) is one of the best studied gene silencing complexes (responsible for H3K27me3 deposition) and has been implicated—usually

through the actions and/or genetic disruption of one or more of its components—in a wide variety of higher phenotypes in mammals. In the cardiovascular system, the Ezh subunits 1 and 2 were shown to be differentially involved in cardiac development and regeneration: both were necessary for normal development, with Ezh1 but not 2 being required for neonatal heart regeneration and with Ezh1 but not 2 being capable, via overexpression, of promoting regeneration in the hearts of mice aged outside the established neonatal regenerative period,[79] suggesting that features of myocyte proliferative/regenerative plasticity may be revived through histone modifying enzymes. Ezh2 stabilizes forming blood vessels in the mouse embryo[80] and pharmacologic inhibition of Ezh2 (and some H3K27me3 target loci) improved outcomes in hind limb ischemia.[81] Some naturally occurring compounds target histone modulating enzymes (and non-histone lysine residue-containing proteins) and have substantial *in vivo* benefit in conditions such as cancer.[82, 83] Since many of these compounds are present in diets shown epidemiologically to promote cardiovascular health (such as cruciferous vegetables), part of this effect may be through actions to promote—at the subcellular level and across organs—favorable epigenomic health.

While the histone isoforms and even specific residues targeted by individual enzymes have often been worked out in reconstituted systems *in vitro*, such information is almost universally lacking from such studies in animal models of the cardiovascular system (this observation is also true, incidentally, for most studies of acetylating/deacetylating and methylating/demethylating enzymes in non-cardiovascular systems when examined at the organ level in animal models). It is also important to note that many of these histone-modifying enzymes target non-histone substrates that have

no direct relationship to gene expression or epigenetics but yet exhibit powerful effects on complex organ level phenotypes (cardiovascular *e.g.* Smyd on titin[70], HDACs on myofilaments[84], and HDAC family members of the Sirtuin class which regulate many substrates in mitochondria and cytosol[56]).

This leads to a couple interesting questions about chromatin modifying enzymes in the cardiovascular system: First, what are the principles that allow for coordination of the various writers, readers and erasers in the given cell type at any time (by this it is meant: how do the histone modifiers themselves get turned on or off, up or down, and when turned on, how do they compete for influence over gene expression in a reproducible manner?). And second, how do the cadre of expressed-at-any-given-time enzymes decide which nucleosomes to modify (*i.e.* how is targeting accomplished, since most histone modifying enzymes do not have DNA sequence targeting motifs)? One approach to answer these questions would be to identify intermediate indices of epigenomic function like accessibility and structure, designing interventions that modulate these indices.

ATP-dependent remodeling of chromatin: inducing and participating in cardiovascular diseases

ATP-dependent chromatin remodeling enzyme complexes use the energy from ATP to translocate DNA through the nucleosome. That is, the ATP-dependent chromatin remodelers reposition nucleosomes along the genome according, in part, to cues harbored in the spectrum of histone tail PTMs, thereby enabling fundamental genomic processes like transcription, nucleosome assembly/disassembly, mitosis, meiosis and

chromosome segregation. One of the most studied of these complexes is the SWItch/Sucrose Non-Fermentable (SWI/SNF) complex, originally identified in yeast (and known to have >10 protein components) and its mammalian cousin the BAF (brahma-associated factor) complex (itself composed of >10 protein components, some of which exhibit tissue specific expression). Models for the actions of these remodelers are informed by protein crystallography studies, *in vitro* biochemical assays and ChIP-seq-based genome wide measurements and are thus highly developed and can help to explain the observed dynamism of chromatin in development, between cell types and in stimulus-response.[85-87] Because presence of bivalent marks at a given locus are necessary but not sufficient to specify a bivalent locus, recent studies have focused on evaluating the role of ATP-dependent chromatin remodelers in this process.[88] Studies in the cardiovascular system have examined the role of these complexes in tissue level phenotypes by genetic manipulation of the ATPase subunits of the BAF complex, Brm (Brahma) or Brg1 (Brahma-related gene 1, a.k.a. Smarca4). For example, genetic disruption of either of these molecules alone had no effect on retinal angiogenesis in neonates, exercise-induced angiogenesis in adult skeletal muscle, or tumor angiogenesis, whereas mice with disruption of both Brm and Brg1 after birth exhibited fatal vascular malfunction in the heart and gut during the early postnatal period[89] (similar context dependent functional redundancy, and lack thereof, was observed between Brm and Brg1 in the vascular endothelium, wherein disruption of both proteins in endothelial cells was required to observe tissue level defects[90]). Brg1 has been shown to be involved in zebrafish myocyte proliferation and cardiac regeneration,[91] mesoderm, and hence cardiomyocyte, differentiation in cell culture, in part by modulating enhancer activity,[92]

whereas both Brg1 and Smarca3 (a.k.a. BAF60c, another BAF complex member) are required for normal heart development in mouse[93-95] (incidentally, Brg1 was found to be down-regulated in adult murine hearts and re-expressed, concomitant with the myosin heavy chain isoform switch [alpha to beta] associated with cardiac pathology; blocking Brg1 upregulation in the adult prevented this molecular event and attenuated hypertrophy[94]). Early formation of vasculature and erythropoiesis in mouse is dependent on Brg1 but not Brm in hematopoietic and endothelial cells,[96] implying functional distinction between complexes seeded with these different ATPases during cardiovascular development (a similar conclusion was made from genetic disruption in smooth muscle cells[97]), a functional involvement that may extend into adulthood in the setting of endothelial injury and presumably disease.[98]

Multifunctional role of DNA methylation in chromatin biology and cardiovascular phenotypes

DNA methylation is dynamic during vertebrate development, where it reinforces cell fate decisions and controls imprinting, or the dependence of gene/protein expression (and associated phenotypes) on whether a given version of a gene is expressed from maternal or paternal allele. Unlike histone modifications, which can occur on any number of different amino acids apparently without heed to locale (that is, without clear DNA consensus motifs), DNA methylation targets a single residue, cytosine, usually in a single context (that is, when followed by a guanine, so called CpG dinucleotides; recent evidence suggests, however, that this too may be an oversimplification as non-CpG DNA methylation, so-called CpH methylation [where the H connotes A, T or C] occurs in some

cells,[99] such as neurons [where it may account for 25% of methylation], although this has only begun to be explored in the cardiovascular system[100, 101]). Also contrasting with the plethora of proteins controlling histone PTM, DNA methylation is directly added or removed by a narrow suite of enzymes: maintenance (DNA methyltransferase 1, DNMT1) and *de novo* (DNMT3a and DNMT3b; DNMT2 modifies RNA and DNMTL is a catalytically inactive regulatory component of the methylation machinery) methyltransferases which establish methylation patterns after mitosis and replication, and alter the pattern of methylation during organismal development and disease, respectively. Demethylation of DNA occurs in part via non-enzymatic means during replication, as well as during normal and pathological conditions in non-dividing cells. Conversion of 5-methyl-cytosine (5mC) to 5-hydroxymethylcytosine (5-hmC) is catalyzed by the ten eleven translocation methylcytosine dioxygenase 1 (TET1) family of enzymes—an active, selective process. 5hmC is a less stable modification, prone to non-enzymatic conversion to (unmodified) cytosine and has thus been proposed as a molecular beacon of genes switching from off to on. DNA methylation patterns are erased and reestablished transgenerationally, but this process appears to involve faithful perpetuation of methylation marks along the genetic lineage, *i.e.* from parent to progeny (further evidence of this phenomenon can be seen in inbred mouse strains, whose DNA methylation landscapes are epigenetically preserved within a genetic lineage, whilst being stably distinct between lineages[102, 103]).

Bisulfite sequencing is the method for unequivocal determination of methylation status (note: all methods for DNA methylation analysis are in symbiosis with suites of informatics tools[104]). Broadly construed, DNA methylation in CpG islands (regions of genome with high frequencies of the dinucleotide, which incline towards promoters) and

shores (areas around said islands) tends to be associated with gene silencing.[105] Conversely, the bodies of mRNA-encoding genes tend to be methylated, without an established correlation allowing prediction of expression. Other methods for large scale analysis of DNA methylation include methylation immunoprecipitation (which has been used to identify methylation dependent regulation of atherosclerotic risk in humans[106]) and DNA methylation arrays (notably the Illumina 450 chips), the latter of which has been extensively deployed in humans to characterize methylation patterns associated with a host of pathophysiological conditions including cancer,[107] high blood pressure,[108] body mass index and obesity,[109, 110] atrial fibrillation,[111] inflammation,[112] and death.[113] Perhaps because of cost and technical demands, reduced representational bisulfite sequencing (or much more expensive whole genome bisulfite sequencing) has been applied to a smaller list of diseases. Such data from mice however show that DNA methylation plays a powerful role in heritable differences in response to metabolic syndrome[102] and may contribute to catecholamine induced cardiac pathology.[103] DNA methylation and/or hydroxymethylation abnormalities have been found in animal[114, 115] and human[116, 117] heart failure, associated with changes in expression of pathologic genes. Work from mouse cardiomyocytes suggests that DNA methylation largely obeys chromatin structural features of A/B compartmentalization (itself defined based on gene density, histone marks and other features of open chromatin; see section below on chromatin structure), wherein dynamics of DNA methylation during lineage commitment are enriched in A (active) compartments and genetic disruption of DNA methylation (via DNMT3a and 3b knockout) does not alter compartmentalization.[101] This observation supports a passive relationship between DNA methylation and chromatin structure, at least in the formation phase.

What is the import that methylation patterns are associated with complex human phenotypes? One method through which DNA methylation has its molecular effects is to reinforce prevailing chromatin landscapes by preventing accessibility and facilitating compaction (in promoters, as mentioned above, and in X chromosome inactivation, where it conspires with histone H3K27me3 to silence expression of one of the two X chromosomes in females[118, 119]); another is to favor relaxing of chromatin and transcription (as in gene bodies). These opposing effects must require the intervention of discriminating factor(s), perhaps including methyl-CpG binding proteins, but this field currently wants for established rules and actors. One investigation[120] demonstrated that MeCP2, a methyl-CpG binding protein, is reversibly downregulated in a mouse model of pressure overload (when the aortic banding was removed, MeCP2 expression was restored; similar observations in patients with LVADs suggest this process may be operative in humans). These findings suggest that the actions of DNA methylation may be modulated in the diseased heart at multiple levels, including methylation, demethylation and reading of methylation. Another mode of action is through *trans* effects, whereby a methylation event can regulate the expression of a gene in a distal region of the genome (*i.e.* far away from the actual CpG in question). Studies from human cardiac development reveal an enrichment of regulatory elements, including DNA methylation sites, in regions of genetic variation associated with heart disease,[121] supporting a molecular link between chromatin regulation and genetic variation in the context of pathological phenotypes. By “regulate,” it is meant here that the methylation event is shown to correlate—with genome-wide statistical significance—with the expression of a gene, a methylation quantitative trait locus. This regulation may take the

form of enhancer element formation/modification (discussed below) or other as-yet uncharacterized chromatin structure effect.

The prevalence of cancer[122] and congenital heart disease[123] in humans is associated with mutations in genes encoding proteins that modify chromatin, such as histone modifying enzymes (writers, erasers and readers). Furthermore, these complex diseases are often associated with global changes in DNA methylation. In some cases, the genetic or epigenetic lesion occurs in a gene whose aberrant function can exert a dominant role in disease pathogenesis. In other cases, these epigenomic changes may instead be general hallmarks of perturbed cellular function, whereby the normal parameter space for gene expression is expanded, facilitating dysfunction of multiple cellular processes. For many observations on histone PTMs and DNA methylation in cardiovascular disease, the *train or snow leopard* question remains unanswered.

Role of noncoding RNAs in chromatin function and cardiovascular physiology

It is now appreciated that most of the genome is transcribed, if only a small portion of that transcriptome encodes mRNA destined for translation, with intriguing differences in this non-coding transcriptome across cell types and following pathological insult. Noncoding RNA biology is a specialized discipline unto itself, with new species of RNA—ascribed really cool and sometimes bizarre functions—identified seemingly endlessly, and will not be extensively reviewed herein (excellent reviews on the roles of various noncoding RNAs in cardiovascular biology have emerged[124-127]). Of particular interest to chromatin biology, however, is the concept that long noncoding RNAs (lncRNAs) may participate in gene regulation by modulating chromatin structure.

lncRNAs have been proposed as a potential mechanism for how different chromatin marks are deposited—specifically and reproducibly—across the genome. One of the best-studied lncRNAs, a general definition of which is an RNA greater than 200 nucleotides with no discernable open reading frames (an exception to this being the presence in some lncRNAs of ORFs which have been shown to produce micropeptides that go on to regulate key intracellular processes in cardiovascular cells[128, 129]), is *Xist*, which is centrally involved in X chromosome inactivation. *Xist* is transcribed from and acts in *cis* to silence the X chromosome through a process that recruits, via direct binding of *Xist* to the proteins, the PRC2 complex and YY1. A depositor of histone H3K37me3 silencing marks and transcriptional repressor, respectively, these proteins in turn compact the X chromosome and prevent further transcription.[119, 130] This model—lncRNA binding to a specific region of chromatin and recruiting histone modifying enzymes—is appealing, because it solves the problem of DNA sequence recognition, of which many histone modifiers are incapable. Another well-characterized lncRNA that binds PRC2 subunits is *Hotair*, involved in gene silencing in mammals, shown to regulate chromatin in *trans* outside of the context of X inactivation.[131] These studies led to a gold rush on PRC2-interacting lncRNAs, which have been estimated to range in number from hundreds to thousands in mice[132] and humans.[133] The general properties, if they exist, through which these lncRNAs couple PRC2 to chromatin are the focus of continued investigation.[134, 135] It may be that the genes for these molecules are distributed across the genome and the lncRNAs in turn all act in a local manner to recruit and modulate chromatin machinery to a given gene expression environment.[136] Yet there are clear limitations were the cell to attempt to repeat this process with other lncRNAs: physiological transcriptional profiles in

adult cells do not involve turning on or off entire chromosomes, with genes temporally co-regulated often residing on different chromosomes (each of which, in this model, would require its own lncRNA, although 3D chromatin environments may allow transcription factories to form bringing multiple mRNA coding genes into a neighborhood governed by a single lncRNA) and beset by numerous histone modifications. Reflecting this fact, the spectrum of lncRNA functions has expanded[119, 137] to include actions in *trans* (*i.e.* targeting other chromosomes) as well as *cis* to enhance transcription, to block it, to scaffold chromatin interactions, and to aggregate microRNAs (thereby making them unavailable to regulate mRNAs).

Initial investigations of lncRNAs in the heart revealed involvement in developmental growth and maturation. *Fendrr* binds both PRC2 and the activating complex Trithorax group/MLL in mesoderm, its depletion leading to impaired cardiac and chest wall development.[138] *Braveheart*, another mesoderm associated lncRNA, binds the Suz12 subunit of PRC2 and is required for proper differentiation of embryonic stem cells into cardiac precursors.[139] Also involved in cardiac development is the lncRNA *Upperhand* that regulates the *Hand 2* locus in *cis* by facilitating chromatin modifications (super enhancer maintenance) and RNA pol II elongation.[140] Other lncRNAs have been discovered to play a role in disease-associated gene regulation. *Chaer* binds the Ezh2 subunit of PRC2 and its genetic manipulation leads to alteration in H3K27me3 levels around pathologic genes and cardiac hypertrophy in the mouse.[141] An antisense transcript in the β-MHC locus was found to associate with that locus in a manner independent of the PRC subunit EZH2 in the setting of pressure overload.[142] Interestingly, that same paper showed the EZH2 interaction with chromatin was regulated by the

noncoding RNA pri-miR-208b, hinting at a broader role for noncoding RNAs in regulating chromatin.[127] Some lncRNAs appear to exert their effects on gene regulation through interaction with chromatin remodeling complexes, as is the case for *Mantis*, a lncRNA discovered in macaque and shown to regulate endothelial angiogenesis in a manner involving interaction with BRG1.[143] Similarly the cardiac-specific lncRNA *Myheart* binds and inhibits the actions of BRG1, thereby regulating expression of myosin heavy chain expression, along with other genes. *Myheart* is downregulated by pressure overload stress and its transgenic restoration protects against overload induced hypertrophy.[144] Interestingly, lncRNAs like *Malat1* in vascular tissues[145] and *Chast* in cardiac tissues[146] are nuclear localized and regulate expression of nearby genes, although it remains to be tested whether they accomplish these actions through recruitment of chromatin complexes. In the field of cholesterol metabolism, two recent lncRNAs have been discovered that exert powerful effects of lipid levels and atherosclerosis *in vivo*: *LeXis*,[147] expressed in the liver, directly controls genes involved in cholesterol biosynthesis, consequently modulating plasma cholesterol levels, and *MeXis*,[148] expressed in macrophages, regulates genes involved in cholesterol efflux (both lncRNAs were found to operate through chromatin based on subcellular localization, accessibility assays and transcriptional regulation).

Discovery analyses in a mouse model of pressure overload revealed the expression profile of cardiac lncRNAs,[149] determining their extent of enrichment in this tissue when compared to tissue of distinct developmental origin (liver and skin) and determining changes between embryonic, adult and diseased non-coding transcriptomes (nota bene: only a few of the developmentally silenced lncRNAs were re-expressed with

disease, in contrast to the “fetal gene program”[150] documented for mRNAs). Likewise in myocardial infarction, the recovery/injury/remodeling period was found, in mice, to be associated with changes in lncRNA expression (incidentally, the lncRNAs were also found to reside near chromatin marks associated with transcriptional enhancement), some of which (the lncRNAs) were subsequently shown to modulate expression of mRNA-encoding genes known to participate in basic cardiac function.[151] Studies from humans have charted differences in lncRNA expression between fetal and adult cardiomyocytes, linking their expression with known enhancer marks associated with protein-coding RNA transcription (*e.g.* H3K4 methylation).[152]

What is known about lncRNAs in the cardiovascular system is that they can be cell type specific, often lack extensive sequence conservation across species (although they may be conserved at the level of secondary structure), can regulate transcription (probably mostly in *cis*) and can correlate with histone PTMs, binding some of the histone modifying complexes, PRC2 in particular. It is unknown to what extent lncRNAs can act at a distance (beyond, say, a few kilobases from their own site of transcription), the role of chromatin structure to coordinate such actions (evidence from X inactivation suggests that local chromatin environment—rather than DNA consensus motifs—facilitates *Xist* binding, PRC recruitment and inactivating activity[130]), if they bind directly to chromatin and/or DNA (perhaps involving triple helix formation[153, 154]) and whether they are sufficient to coordinate the locus specific activities of chromatin modifiers through a model in which multiple lncRNA genes, by virtue of their evolutionary population at key sites across the genome, establish local neighborhoods of regulation at which they recruit—or repel, according to wont—histone modifiers and polymerase machinery.

Cardiovascular development- and disease-associated enhancer elements

Enhancers are regions of DNA that promote the transcription of other regions of DNA.[155-157] A contemporary synthesis on how this works: specific histone PTMs (*e.g.* histone H3K4me1 and H3K27ac; some histone isoforms, such as H3.3 and H2az, contribute to enhancer activity; enhancers are now thus commonly identified by genome-wide ChIP-seq experiments) decorate regions of DNA that need not be—although may be (see below)—themselves transcribed, which in turn recruit binding of enhancer associated proteins (*e.g.* lineage relevant transcription factors, RNA pol II and co-activator proteins, such as p300 and Mediator) and interact in three dimensions with the genes whose expression they enhance. This region of DNA, the appropriated demarcated histones and any associated proteins, together constitute the enhancer which is often "validated" as such by showing that either (a) its genetic disruption interferes with expression of its target gene and/or (b) that the enhancer DNA sequence can drive developmental and lineage appropriate transcription through a cell- or organism-based reporter assay. In the absence of chromatin conformation data, enhancers are usually assumed (and tested) to regulate the nearest downstream gene. Somewhat counterintuitively, then, enhancers tend to reside in areas of relative nucleosome depletion (not, strictly speaking, in areas devoid of nucleosomes), such that enhancers can be identified by open chromatin assays (*e.g.* DNAse I hypersensitivity or ATAC) followed by DNA sequencing. A subgroup of enhancers, called super enhancers,[158] has been classified based on the observation that the aforementioned enhancer features at times occur multiple times in close proximity to each other. Super enhancers can exhibit

augmented transcriptional activation potential and thus may represent a distinct structural property of cell type-specific chromatin.[159] Further specification of enhancer behavior includes delineation of poised (those ready for promoting transcription of their targets) versus active (those actually so promoting) enhancers, which can be distinguished by the presence of silencing histone marks (and the enzymes that deposit them) at poised enhancers and their absence (concomitant with the presence of greater levels of RNA pol II) at active enhancers.[155-157]

It has more recently become apparent that some enhancers may themselves be transcribed[160] and may thus operate in the RNA form. It could be that this transcription is a goal-directed process in the normal way we think about RNA *doing things* in the cell: the enhancer RNAs, (eRNAs) may have gene regulatory or other functions. It may also be that the eRNA synthesis is a by-product of enhancer DNA in close apposition to churning transcription factories and serves no subsequent end and/or that its very transcription serves the end of keeping a transcription factory churning and poised for ready enlistment in production of other RNAs that do serve subsequent ends (as RNAs). This fascinating concept of chromatin biology is an active area of investigation.[161] Studies have begun to emerge examining the role of enhancer transcription in select cardiovascular processes such as cardiac conduction[162] and endothelial cell stress response.[163] Active endothelial cell enhancers, defined by H3K4me2 and H3K27ac binding (plus some eRNA transcription), exhibited altered transcription factor binding in human aortic endothelial cells following exposure to oxidized phospholipids.[163] Intriguingly, SNPs associated with cardiovascular disease were over represented in these enhancers, suggesting a molecular scale explanation for how the former influences

transcription and phenotype, a property that may be a common feature of enhancers across cell types and species.[164]

p300 occupancy has been used to identify enhancers in the developing mouse heart embryonic day 11.5, many of which were found to exhibit tissue specific activity.[165] A similar approach was used to characterize enhancers in fetal and adult human heart tissue (of note, 48% of the enhancers were the same in fetal and adult human hearts; when comparing fetal mouse to fetal human, the overlap was 21%),[166] revealing functional elements that may participate in human cardiac gene regulation. Angiotensin II-induced vascular growth, a key component of atherosclerosis, was found to proceed via dynamic utilization of super enhancers in human cells through a process that involves complex interplay amongst noncoding RNAs, chromatin readers and transcriptional machinery.[167] A theme of developmental processes being redeployed—not wholesale, but in a selective manner—in the disease setting may also play out for enhancers: regulatory elements marked by H3K27ac are specified in part by the master cardiac transcription factor GATA4 during development and some of these regions, devoid of GATA4 in healthy adult heart, are revisited by the protein upon pathologic stimulation, contributing to disease-associated gene expression.[168] Pursuing this concept more directly in a complementary model, it has also been shown that cardiac enhancers undergo altered regulation by disease-associated transcription factors following pathologic stress.[169] The histone modification reader BRD4 binds super enhancers that are associated with cardiac disease genes. Interestingly, this process is finely tuned to differentially modulate association of BRD4 with these disease genes while leaving

housekeeping genes unaffected, a process controlled in part by miRNA-dependent titration of BRD4 levels.[170]

Unexpected cell type-specific functions of chromatin

The rapidly dividing phenotype of cancer has allowed researchers in this field to identify epigenetic clones:[171] lineages of cells outwardly genetically identical that differ based on semi-stable, transmissible chromatin features. All cardiac myocytes and vascular smooth muscle cells are not the same, which means that although these cells do not proliferate and differentiate like cancerous cells do, it is reasonable to hypothesize that developmentally endowed epigenetic clones exist and contribute to organ level phenotypes in the adult. Indeed, distinct clonal populations (arising from a common progenitor in development, rather than from a resident adult stem cell) of cardiac cells contribute to different anatomical and functional features of the adult organ[172-177] (recent single cell studies have revealed these distinct myocyte populations to indeed exhibit distinct transcriptomes[178, 179])—epigenetic dissection of these populations may well reveal epigenetic clonality to be an underlying process contributing to this observation. The chromatin accessibility assay ATAC-seq, which reveals areas of open chromatin, has been applied in a single cell format to a lymphoblastoid cell line, identifying subpopulations of cells based on chromatin accessibility.[180] That such variability exists in post-mitotic, healthy adult cells in the cardiovascular system remains to be demonstrated, but the observation of transcriptome variability in these cells, and the presence of chromatin accessibility variability in cells otherwise phenotypically similar, makes such a conjecture not unwarranted.

Chromatin can act like a stress sensor complex, wherein there is no single factor controlling changes in disease associated gene expression. Some investigators have described excitation-transcription coupling, with the term specifically applied to local calcium signals around the nucleus (as distinguished from global calcium transients involved in myofilament contraction) inducing local CaMKII activation and HDAC mobilization.[181] What if this observation is evidence of a more generalized, myocyte specific sensory apparatus on chromatin, that detects local calcium signaling, such as that involved in pathologic gene activation, from calcium involved in contraction and nonetheless critical to influence gene expression (*e.g.* sustained faster heart rates require greater turnover of proteins and thus transcripts)? Various pathological cell states, including cardiovascular disease and cancer, have been characterized by global changes (*e.g.* that revealed by a total cell lysate western blot or genome-wide ChIP-seq signal, for example) in histone modification. One possible reason for this unexpected observation was found to include regulation of cellular acidity:[182] global histone acetylation responds to perturbation of cellular pH (lower pH leads to less histone acetylation) and cells respond to modulation of histone acetylation by modulating pH, a sort of acetate capacitance system on chromatin to attenuate large swings in cellular acidity. Combined metabolomic and proteomic studies reveal that abundance of short chain acyl-CoA donors directly, although not indiscriminately, influences the modification of histone tails in human cells in culture.[183] Indeed metabolic sensing by chromatin has been increasingly recognized to underpin cardiovascular physiology and disease.[76]

Aberrations in nuclear rigidity and structural integrity are associated with diseases like cancer and progeria, some of which are driven by so called laminopathies, arising

from malfunction of nuclear lamina proteins. Cardiomyopathies resulting from mutations in lamin A/C are one of the best studied group of genetic diseases in clinical cardiology, and have led to clinical trials, although in this context the effects on chromatin structure (or vice versa) are unclear: aberrant nuclear morphology, the blebbing of the nuclear membrane due to impaired laminar network architecture, is a hallmark of these diseases.[184] Association of chromatin with the lamina appears to be essential for nuclear structure and disrupting this interaction has detrimental effects on nuclear integrity,[185] particularly in cells subject to mechanical force. These actions are coupled to the mechanisms known to regulate chromatin's role in gene expression, as supported by the observation that histone PTM influences nuclear rigidity and membrane integrity.[186]

An unexpected non-nuclear, signaling behavior of not just chromatin modifying proteins but actual intact multimolecular slabs of chromatin has been observed in cancer[187]: cytoplasmic chromatin fragments—evaginated nuclear membrane containing DNA and nucleosomes decorated with heterochromatin marks—can induce inflammation and cell death through cytoplasmic signaling and circle-back transcriptional regulation. An even weirder story: rod and cone cells are terminally differentiated, specialized components of the retina, the light sensitive component of the vertebrate eye. Evolution has hijacked chromatin in rods (but not cones) to serve the transcriptionally unrelated function of focusing light in the retinas of nocturnal but not diurnal mammals (that is, the organization of DNA in the nucleus forms a physical lens),[188] thereby providing a meta-function in service of that specific cell's raison d'être.

Different structural units of chromatin establish distinct transcriptional environments. It can be helpful to think of chromatin itself as a transcription factor, or

transcriptional processor. Transcription does not happen willy-nilly throughout the nucleus, but rather is localized to transcription factories,[189] or areas designated to different forms of transcription, such as rRNAs and house-keeping sorts of protein-coding mRNAs, separated from stimulus responsive genes and furthermore from transcriptionally silent regions. This happens on a nuclear scale as reflected by the observation that transcription tends to happen toward the center of the nucleus whereas the periphery is an area of gene silencing. Another type of transcription factor-like activity is sub-chromosomal, in the form of chromatin looping, the formation of short- and long-range interactions to facilitate gene activation (*i.e.* enhancer elements) or repression (*i.e.* insulators or boundary elements). Chromatin capture data has revealed that long range looping within the epigenome is dynamic and can bring together transcription start and end sites in three dimensions, perhaps to facilitate efficient cycling of machinery like polymerases and transcription factors (Hi-C data supports this concept for a cohort of genes, Figure 1-3 and Table 1-2). This principle has been supported with ChIP-seq data from rat hearts,[190] in which different transcriptional activation profiles (pause-release and *de novo* recruitment) have been described in the setting of pressure overload hypertrophy, along with accumulation of RNA pol II in transcription end site, perhaps reflective of gene looping.

Structure-function features of chromatin and implications for cardiovascular gene regulation

Based on insights from chromatin capture and other epigenomic techniques, the organization of the epigenome is thought to involve key structural intermediates (Figure

1-4). What is the evidence that chromatin is inherently ordered above the level of the nucleosome (where data exists to the atomic—that is several angstrom—level[2]) and below the level of the chromosome (where chromosome painting, closer to the scale of micrometers, demonstrates compartmentalization[191])? The goal of chromatin structural studies is to determine: what are the structural features between these scales and at what scale(s) are structural features functionally important? Next consider the pattern of chromatin observed in various cells of the cardiovascular lineage with a quotidian method such as DAPI labeling: while the pattern of staining is not random, there is no obvious reproducible pattern within a class of cells (and not shared between two classes) to which a functional consequence can be intuited (for a nifty exception, see[188]), in contrast to chromosome patterns in mitosis/meiosis which definitively exhibit such tell-tale architecture. Fluorescence *in situ* hybridization (FISH) experiments clearly demonstrated spatial segregation of chromosomes into territories,[191] whilst not revealing evidence of hierarchical arrangement. Recent higher resolution electron microscopy-based imaging of chromatin shows that its structure, in both interphase and mitotic cells, rarely achieves a scale greater than 24 nanometers in diameter (for reference, the nucleosome diameter is ~11 nm), with distinctions in arrangement between such cells coming from the density of compaction, which the authors interpret to be evidence of an absence of repeating, stable, hierarchical structure.[192] How do these observations hold up in analyses of individual genes and with respect to histone post-translational modifications? In the cardiovascular arena, combination of Dam-ID and LaminB ChIP-seq (to identify loci associated with the nuclear periphery) and FISH was used to demonstrate that differentiation in the myocyte lineage involves precise reorganization of expressed genes

away from the myocyte nuclear membrane, itself found to be decorated with the silencing mark H3K9me2 and to a lesser degree by H3K9me3 (other silencing marks H3K27me2/3 and H4K20me2/3 were not found enriched at the periphery in skeletal myoblasts).[193]

A prediction of a non-hierarchical model of chromatin is that the size of structural elements should be normally distributed. For chromatin interactions detected by Hi-C, for example in cardiac myocytes, this prediction has been shown to be true: the number of interactions plotted per locus follows a normal distribution, where most locus bins (bin size=5kb) have the same number of interactions (~2500) and a small number have very few or very many interactions (see Supplemental Figure 1a in reference[194]). The vast majority of loci interact with a median number of other loci, and no privileged structural behavior can be assigned to the regions with large interactions (as would be a prediction of a scale free or hierarchical topology).

Additional insights from the explosion of chromatin capture techniques to determine endogenous interactions have been informative.[195, 196] These studies have characterized features of chromatin organization that are conserved across species and cell type (note: method development for analysis of chromatin capture data is ongoing and the interpretation evolves with it[197]): topologically associated domains (TADs) are regions of chromatin with privileged local interaction; TAD boundaries, as nominally implied, demarcate regions of the genome where intrachromosomal interactions switch from interacting in one direction (say 5' biased) to the opposite; distinct regions are insulated against expression, in part by chromatins structural proteins and histone modifications; short and long range interactions reproducibly form (*i.e.* the structure is not random). The boundaries represent epigenomic cornerstones, directing interactions of

nearby DNA and proteins in alternating directions, in part through the binding of chromatin structural proteins like CTCF. Cohesin and CTCF knockout animals and cells indicate these proteins are involved in TAD maintenance,[194, 198, 199] but these proteins alone are not the whole story: embryonic stem cells[198] with only 4% normal CTCF protein levels still exhibited TADs and cardiomyocytes[194] with 20% normal CTCF protein levels exhibited sparse, minor changes in TAD boundaries and strength. These studies also show that TAD formation/maintenance and A/B compartmentalization can be decoupled experimentally, as loss of cohesin or CTCF did not affect A/B compartmentalization. “A” compartments have more genes and are defined by having less interactions than would be expected for a given distance (“B” has more), indicating less compact chromatin. Eu- and heterochromatin marks dominate in A and B compartments, respectively. Prevailing evidence suggests that at the level of TADs, chromatin architecture is quite similar between cell types. Sub-TAD interactions, and less abundant—yet reproducibly detected—long-range interactions that span multiple TADs (as well as inter-chromosomal interactions), may be the scales at which cell type specific chromatin interactions are observed.

Chromatin structure is dynamic during the cell cycle. The predominance of longer range interactions present in post-mitotic cells are rapidly lost upon entrance into G1, followed by further depletion throughout S and G2, in favor of shorter range, local interactions. This process abruptly reverses itself, with a return of TADs and long-range interactions following nuclear division.[200] Mitotic chromosomes lack TADs and chromatin neighborhoods, instead exhibiting a uniform, homogenous pattern of hierarchical interactions.[201] A similar observation was made for oocytes in metaphase II: an absence

of TADs and chromatin neighborhoods in these cells persisted in the zygote, with long range chromatin interactions manifesting at the 8-cell and inner cell mass stages.[202] Interestingly, physical segregation of alleles was seen to persist until the 8-cell stage as well,[202] even after the formation of long range chromatin contacts, suggesting that chromatin structure is an emergent property of an allele, can vary between alleles and thus may participate in allelic inheritance.

Probabilistic modeling has been used to reveal 3D organizational principles from Hi-C datasets, the goal here being not a structure *per se*, but a population-based representation of the structural features of the chromosomes as they associate in the nucleus.[203] Using datasets from human lymphoblastoid cells, this approach was used to schematize genome structure, revealing inter-chromosomal surfaces of apposition and detecting new anatomical properties of the nucleus, such as the physical clustering of centromeres of different chromosomes and the anatomical positioning of euchromatin and heterochromatin pockets with respect to other nuclear landmarks. Unlike traditional FISH or chromosome painting, in such an exercise one can know which loci are responsible for a given anatomical feature and in what part of the nucleus this feature tends to occur relative to other feature.

The histone code[17] has become a ready-to-hand tool for chromatin interrogation, shaping how studies are designed and interpreted. However, as discussed elsewhere in this essay, hundreds of histone PTMs are now known to exist. Also, the histone code and related ideas[204] lack an underpinning in mathematical rules, a limitation addressed by investigators who have used dry lab approaches to characterize chromatin states[205] or rules of nucleosome positioning[206] that reconcile ChIP-seq and chromatin accessibility

data with genome sequence and transcription. Apart from the accordant histone and chromatin binding proteins associated with different flavors of chromatin, how distinct chromatin domains form, in a physical sense, is not completely understood. Recent evidence from *Drosophila* and human chromatin suggests that heterochromatin domains comprised of H3K9me3-marked nucleosomes, heterochromatin protein 1 (HP-1) and DNA can exhibit phase separation behavior, which may be an explanation to link domain-scale and molecular-scale properties of heterochromatin foci which can display both liquid and stable phase properties,[207] a phenomenon supported on a broader scale by contemporaneous studies.[159] The way forward is to integrate structural studies from imaging and sequencing techniques with genome occupancy studies from sequencing techniques to build a new model governed by principles that incorporate all these sets of data.

Integration by the epigenome of genetic susceptibility and environmental risk

Histones were originally identified as inhibitors of transcription. This concept remains a kernel of chromatin theory: heterochromatin increases during differentiation and loss of pluripotency, and some diseases, notably cancer, have been found to be associated with a more euchromatic environment. Because they change concomitant with gene expression and phenotype, chromatin modifications are *ispo facto* taken as responsible for the unidirectional progression of cell fate commitment in the cardiovascular system. Another hypothesis cached therein is that histone PTM and other chromatin marks stabilize cell identity. Regarding this conclusion, here is a premise that should be rejected: identification of a chromatin modifying enzyme in the heart or

vasculature whose genetic manipulation impairs or reverses developmental state is a necessary and sufficient condition to prove a role for chromatin in deciding and/or stabilizing cell fate. Here is another such premise of tenuous utility: if chromatin modifications stabilize cell phenotype, then reprogramming strategies that restore pluripotency (*e.g.* iPS) or directly convert one cell type to another[208, 209] must do so by wholesale reprogramming of chromatin (although these processes do, no doubt about it, reprogram histone post-translational modifications and DNA methylation at cardiac genes[210]). iPS-derived cells coaxed toward a cardiovascular lineage acquire regulatory elements (*i.e.* histone post-translational modifications on regulatory elements nearby lineage appropriate genes) reminiscent of their endogenous counterparts,[211] and yet studies from non-cardiovascular tissues have shown that iPS-derived cells retain some epigenetic memories from their cells of origin, which, perhaps not surprisingly, is also the case for cardiovascular cells derived in cell culture from developmental precursors (*e.g.* DNA methylation).[212] Cardiac cells exhibiting progenitor-like behavior isolated from adult hearts indeed exhibit, commensurate with transcriptome changes, DNA methylation changes in genes associated with the mature cardiomyocyte lineage vis-à-vis adult cardiomyocytes lacking such progenitor-like behavior.[213]

Cell culture studies of distinct stages of cardiac lineage commitment explored the changes in chromatin marks associated with this process.[214, 215] General features of heterochromatic mark (H3K27me3) loss were observed around genes that were expressed (genes never expressed in the cardiac lineage, in contrast, retained abundant H3K27me3 through differentiation and never gained activating marks like H3K4me3), and genes that would be expressed in subsequent stages of development were sometimes

(although not always) enriched with H3K4me1 (a so-called 'poised enhancer' mark) prior to acquisition of H3K4me3 and RNA pol II concomitant with expression. If one were so inclined, the following observations may be taken as evidence that chromatin becomes more plastic in the setting of cardiac pathology (see also Figures 1-5 and 1-6): stimulation of neonatal rat ventricular myocytes with isoproterenol leads to decreased density of chromatin as measured by histone H3 immunolabeling and super resolution microscopy;[216] pressure overload hypertrophy is associated with a decrease in total histone H3K9me3 and increase in total H3K4me3 (as detected by western blotting) as well as a decrease in the linker histone H1 to core (measured by H4) ratio;[8] association across genetically variable mouse strains between select chromatin structural proteins HMGB2 and CTCF and cardiac phenotype and the ability of HMGB2 to modulate cardiomyocyte hypertrophy and chromatin accessibility in cardiac myocytes;[217] and loss of CTCF (which induces cardiac dysfunction) or pressure overload hypertrophy is accompanied by a global decrease in genomic interactions detected by Hi-C.[194]

Meta-analysis of heart failure GWAS studies recently uncovered a novel risk allele associated with mortality and located in a noncoding enhancer region.[218] Interestingly, DNA methylation signatures at this locus in blood were correlated with allergic sensitization, potentially hinting at a gene-environment interaction leaving an epigenetic signature. A similar observation of epigenetic risk conferred in a cell type specific manner by marks present across different tissues was recently reported in the context of dilated cardiomyopathy.[219] The durability of these marks in a temporal sense remains an open question. Unequivocal determination of transgenerational inheritance of chromatin features like DNA methylation or histone modifications (discussion of which often

conflating the two distinct concepts of epigenetics and Lamarckian evolution or the inheritance of acquired traits) is tricky[220] and many examples are hotly debated. Epigenome-wide association analyses in liver demonstrate[102] DNA methylation-dependent—and sequence variation *in*dependent—associations with clinical traits important for cardiovascular disease such as insulin level, as well as other 'omics endpoints, providing proof of concept for population level epigenomic regulation of complex disease traits through the actions of DNA methylation variation to control phenotype, presumably through effects on chromatin structure or accessibility. Something ostensibly heritable through cell division or meiosis may masquerade as epigenetic and/or may even be of chromatic origin but may in fact proceed via a genetic means.[221]

It has been increasingly recognized that epigenomic modifications are influenced by various diet and lifestyle factors that affect cardiovascular health (reviewed in detail elsewhere[222]). Maternal smoking, for example, is known to induce widespread DNA methylation differences in newborns, some of which persist into the offspring's adulthood, including in genes known to be associated with smoking-related birth defects,[223] although whether these effects are due to natal exposure remains unknown. There have been limited experimental studies directly testing the role of non-genetic inheritance of cardiovascular risk in animal models (*e.g.* DNA methylation dependent target gene expression in the context of offspring ischemic injury[224]). *In vitro* fertilization experiments in mice using gametes from obese or normal weight parents (note: obesity was induced by high fat diet) and surrogate, normal chow fed, mothers revealed that a propensity for increased body weight can be inherited by non-genetic means.[225] Metabolic gene

regulation has been shown to be a modifiable, and subsequently heritable feature, in that mice fed low protein diets passed hepatic gene expression profiles transgenerationally through the paternal germ line, commensurate with heritable changes in DNA methylation (although to what extent differences in DNA methylation resulting from distinct paternal diets directly control gene expression profiles remains unknown).[226] Genetic variability contributes to chromatin accessibility (measured via FAIRE-seq) in the basal state and following complex metabolic changes, such as those accompanying high fat diet.[227] Human studies in ethnically diverse populations have revealed DNA methylation variation associated with nicotine and alcohol dependence (and the co-dependency between these forms of addiction)[228] although no evidence of inheritance and/or precedence of the phenotypes by the epigenetic features was demonstrated. In addition to the obvious prognostic and diagnostic potential of genomic and epigenomic measurements in the clinical arena (the practical considerations of which are discussed in detail elsewhere[229, 230]), it is noteworthy that tools for reducing epigenomic treatment to practice have begun to emerge (Figure 1-7). Modification of the CRISPR/Cas9 gene targeting system, involving an inactive Cas9 nuclease (so-called dCas9) fused to DNMT3a or Tet1 and combined with guide RNAs to localize the complex, can induce altered DNA methylation of specific loci in somatic cells *in vivo* commensurate with desired changes in gene expression[231] (techniques for remodeling chromatin loops with designer CRISPR tools have also emerged[232]). Such approaches may enable targeting of entire transcriptomes, rather than individual molecules, in a gene therapy workflow that at once provides both specificity and temporal tuning.

Genomic and epigenomic technologies are providing high resolution descriptions of susceptibilities and etiologies of cardiovascular diseases—that is, we now have the ability to acquire a far greater number of data points on patient populations, enabling identification of new biology but also, in the clinical setting, better stratification.[230] The other 'omics technologies (such as proteomics, lipidomics and metabolomics) that are now experimentally mature can, when applied to chromatin, provide greater still molecular detail regarding how transcriptomes are specified and cell type-specific behavior governed in health and disease. Effective utilization of this wealth of knowledge to promote human health will require innovative thinking: some strategies will involve multi-marker panels (*e.g.* DNA sequence variants measured along with protein or lipid levels to make diagnoses, such as in hypertension) whereas other strategies will use the intermediate endpoints that emerge from the collective actions of multiple classes of molecules as the readout for diagnoses or as target for treatment (*e.g.* targeting chromatin readers or chromatin accessibility, both of which integrate the actions of various signaling processes, protein modifiers, metabolites and RNAs acting in the context of genetic variation).

It occurs to us that another definition of epigenomes would be: the molecular features that make a living creature the same unit tomorrow that it is today. As described in the preceding sections, recent studies have identified the actions of discrete protein components and modifiers of chromatin in cardiovascular health and disease, providing potential targets for therapy. Epigenomic investigations have described chromatin landscapes, providing the data necessary to discover principles for the actions of the protein and RNA modifiers. These epigenomic investigations also enable discovery of

intermediate properties, like chromatin accessibility and structure, which may be traits that contribute to higher level phenotypes. As the integrator of genetics and environment, and the substrate of cellular memory, chromatin features may provide the basis for understanding normal and pathological cardiovascular function.

Acknowledgements: We regret that space limitations have forced us to omit reference to many important papers in the field of cardiovascular epigenomics. We thank Dr. Emma Monte for critical feedback on this manuscript. Work in the Vondriska lab is supported by the NIH, the AHA, and the Cardiovascular Theme in the David Geffen School of Medicine at UCLA.

Author contributions: Concepts: all authors; Text: TMV; Figures, MRG, DJC; Final version: all authors

The authors have no relationships to disclose regarding the content of this paper.

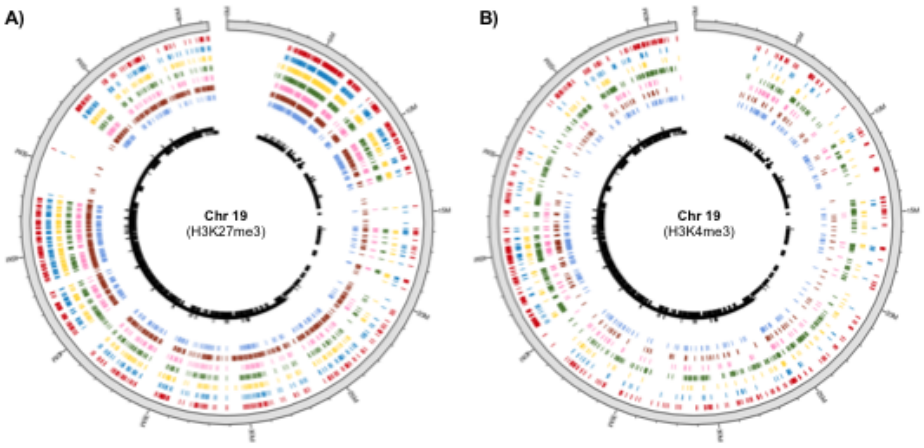


Figure 1-1. Occupancy of histone marks varies between cell types. Circos plot of mouse chr19 shows A) repressive H3K27me3 and B) active H3K4me3 peaks across different tissues. Color-coding of tissues (outside to inside): heart (red), cerebellum (blue), kidney (yellow), liver (green), thymus (orange), testis (dark red), and spleen (light blue). The black track represents mm10 gene density on chromosome 19. Data are from the ENCODE database.

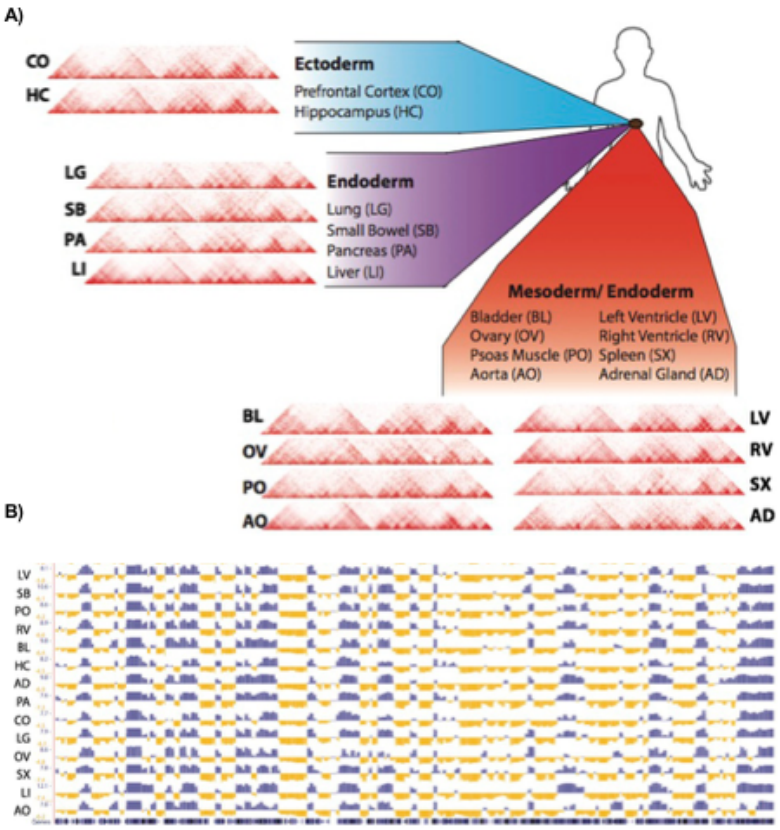


Figure 1-2. Chromatin conformation capture data reveal similarities in chromatin organization between cell types. A) Representative contact matrices from 14 human tissues (A), and their A/B compartmentalization profiles (B), demonstrate similarities at the scale of TADs and compartmentalization. LV, left ventricle; SB, Small Bowel; PO,

Psoas Muscle; RV, Right ventricle; BL, Bladder; HC, Hippocampus; AD, Adrenal Gland; PA, Pancreas; CO, Prefrontal cortex; LG, Lung; OV, Ovary; SX, Spleen; LI, Liver; AO, Aorta. Figure adapted, with permission, from Schmitt et al., Cell Reports (2016).

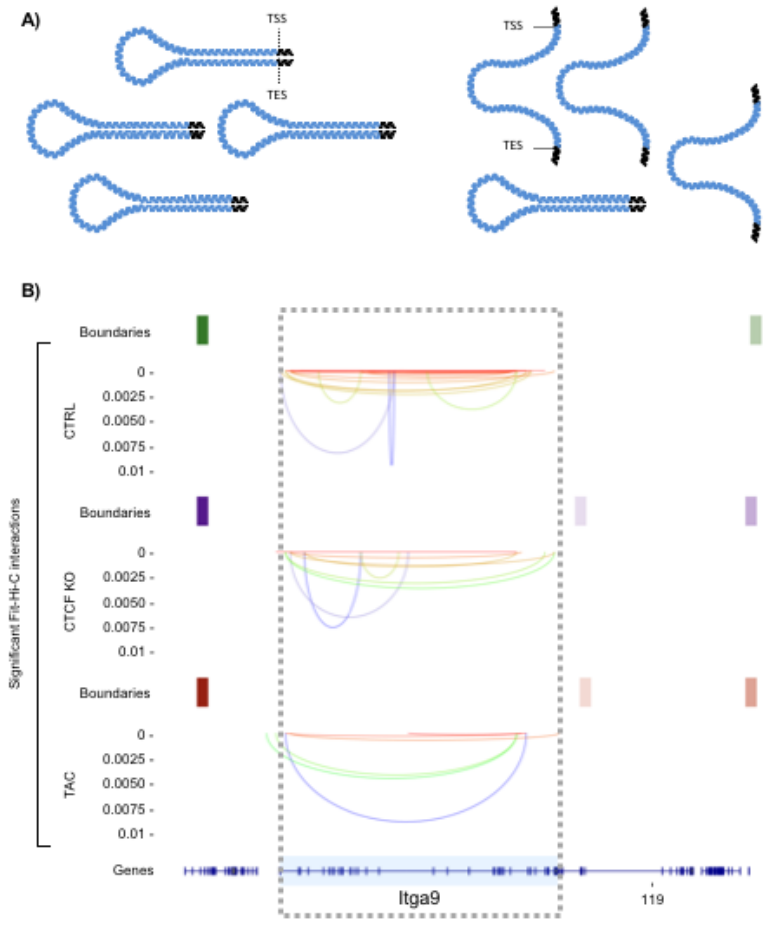


Figure 1-3. Chromatin looping. A) Schematic representation of chromatin looping, in this example between transcription start (TSS) and end (TES) sites of a gene. The model is an interpretation of chromatin capture data (which shows a decrease in interactions during cardiac pathology) and is intended to represent the frequency of a given conformation, not a population effect across cells: left panel shows that under normal

conditions, loops are stably formed; right panel shows that because loops are less stable, they are less frequently captured experimentally. B) An example gene displaying this behavior. Top is control, middle is CTCF knockout and bottom is transverse aortic constriction. Rectangles are TAD boundaries and lines are chromatin interactions detected by Hi-C (Rosa-Garrido et al. Circulation. 2017). Table 1-2 shows a list of genes that undergo the phenomenon described in the figure.

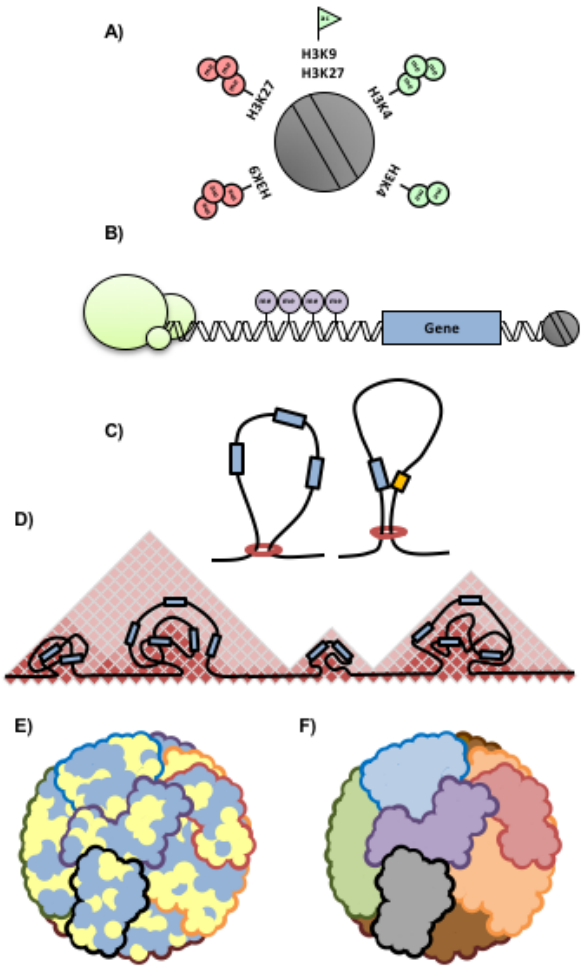


Figure 1-4. Chromatin architectural features. A) The functional unit of chromatin is the nucleosome, which can be decorated by a variety of post-translational modifications tails that modulate accessibility to transcription factors or chromatin modifiers. B) At the gene level, transcription factors or repressors (green circles) confer context-specific regulation

of transcription with varying levels of sequence specificity. DNA methylation (purple circles) typically repress promoter activity of genes, although this phenomenon is associated with expression when found within gene bodies. C) Chromatin looping enables formation of gene expression or silencing neighborhoods, as well as facilitating structural units suitable for higher order packing. D) Topologically associating domains (TADs) are regions of preferential chromatin interactions. E) Hi-C data reveals chromatin compartmentalization into “active” and “inactive”, or “A” and “B” compartments of the genome, respectively (here shown in yellow and blue; note, this is a stylistic interpretation of how A and B compartments might interact, because chromatin capture studies do not reveal actual localization coordinates within the nucleus). F) Chromosome paint experiments have revealed distinct territories that contain entire chromosomes within the nucleus, allowing formation of intra- and inter-chromosomal interactions that may regulate transcription or other tasks of the nucleus.

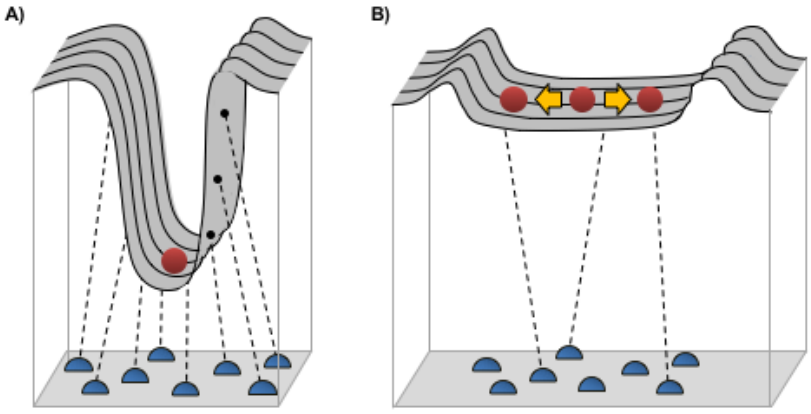


Figure 1-5. Plasticity in epigenomic landscapes may allow for transcriptome reprogramming in disease. Adapting Waddington's concept of, to paraphrase, the chemical tendencies underpinning the epigenetic landscape (The Strategy of the Genes. New York: The Macmillan Company; 1957), the figure depicts how, when some of the chemical tendencies (A; which we now know to be the DNA, proteins and RNA that establish the 3D structure of the epigenome, blue semicircles) are perturbed by

experiment or environment (B), the red ball (which here represents a cell or cell population) can adopt different positions along the energy landscape, becoming sufficiently plastic to enable disease-associated gene expression.

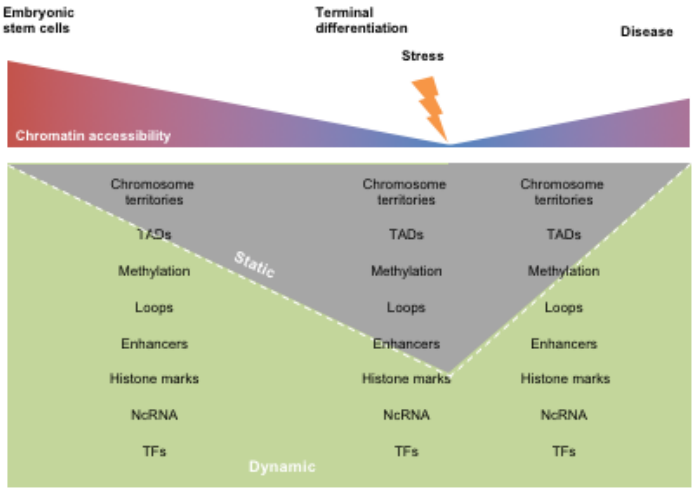


Figure 1-6. Model for epigenomic changes in development and disease. Development is accompanied by changes in chromatin structure and regulation to endow terminally differentiated cells with stable transcriptomes. Disease upsets this balance, transitioning select regions of the genome into more dynamic conformations through effects on chromatin structure, enhancer-gene looping and alterations in histone modification, DNA methylation and other factors. This model is based on findings reviewed in the current paper and adapted from Rosa-Garrido et al., Circulation 2017.

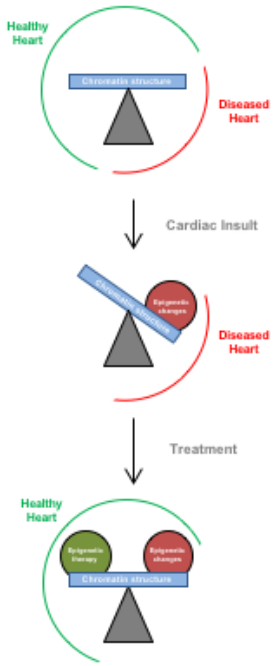


Figure 1-7. Cardiac epigenetic therapies. Changes in chromatin structure during cardiovascular pathologies alter gene expression and thereby phenotype. Epigenetic therapies (examples include HDAC inhibition, BET inhibition, chromatin structural protein modulation, and chromatin loop or DNA methylation targeting by CRISPR/Cas9 tools) could be designed to reverse these changes by targeting intermediate chromatin features like accessibility and structure.

Family	Gene	Manipulation	Model	Phenotype	Molecular Targets	Reference (PMID)
Histone Methyltransferases	Ehmt2	Knockout	Mouse	Required for cardiomyocyte homeostasis Its depletion causes cardiac hypertrophy	Required for the maintenance of heterochromatin that silence developmental genes in the adult heart	28778944
	Ezh1	Knockout	Mouse	Normal embryo. Ezh1 and 2 double KO causes lethal heart malformations, including hypertrabeculation, compact myocardial hypoplasia, and ventricular septal defects. Ezh1 improves heart function and stimulated cardiac myocyte proliferation	Development and cell proliferation genes	28512107
	Ezh2	Cardiac specific knockout	Mouse	Normal embryo. Ezh1 and 2 double KO causes lethal heart malformations, including hypertrabeculation, compact myocardial hypoplasia, and ventricular septal defects	Development and cell proliferation genes	28512107
				Perinatal lethal; Double outlet right ventricle; Underdeveloped endocardial cushion due to decreased proliferation and increased apoptosis; Impaired septal formation, hypertrabeculation; Hypoplasia	Upregulation of negative cell cycle regulators (e.g. Ink4a, Ink4b), non-cardiac transcription factors (Pax6, Six1, Isl1), atrial and conduction gene expression in ventricles (Hcn4, Mlc2a); Downregulation of Tbx2, Hey2 transcription factors	22158708, 22312437
		Knockout		Embryonic lethality with compromised vascular integrity and increased extracellular matrix degradation	Mmp9 and its activators Fosl1 and Klf5	25359725
	Kmtd2	Knockout	Mouse	Embryonic lethality Conditional deletion at E10.5 shows developmental delay, pericardial edema and a linear heart tube	Essential for regulating cardiac gene expression during heart development via H3K4 di-methylation	26932671
	Mll	Knockdown	HUVEC Cells	Reduction in endothelial-cell migration and sprouting	HoxA9 and HoxD3	17047146
	Mll2	Knockout	Mouse	Embryonic lethal by E14.5 Apoptotic	-	16540515, 23826075
		Point mutation		Embryonic lethal by E11.5; Impaired looping		
	Setd2	Knockout	C2C12	Defects in myoblast differentiation and significantly reduction of histone 3 phosphorylation leading defects in cell proliferation	Myosin heavy chain (MHC) and Myogenin (MyoG) downregulation	28130125
	Smyd1	Knockdown	Zebrafish	Heterozygotes show no noticeable impairments	Upregulation of heat-shock proteins	16477022, 24068325
				Impaired cardiac function Disrupted sarcomeric organization		
		Knockout	Mouse	Cellular hypertrophy, organ remodeling, and heart failure	Transcriptional repression of Tgfbeta3 and Nppa genes	27663768
	Whsc1	Knockout	Mouse	Lethal by P10 and growth defects Impaired septal formation due to hypoplasia	Upregulation of Nkx2-5 dependent genes (e.g. Pdgfra, Isl1)	19483677
				Heterozygotes display some phenotypes of knockout due to haploinsufficiency		
Histone Demethylases	Jarid2/Jumonji	Knockout	Mouse	Perinatal lethal; Thin ventricular walls; Double-outlet right ventricle; Impaired septal formation, hypertrabeculation; Edema; Increased proliferation	Maintained fetal gene expression (increase in βMHC to αMHC ratio, ANF not repressed); Impaired patterning of cardiac gene expression (e.g. ventricular ANF expression maintained, high Mlc2a expression in ventricles and atria); Increase in Notch signaling	10807864
				Increase in mitotic activity in isolated cardiomyocytes	Upregulation of Cyclin D	15870077
				Perinatal lethal; Thin ventricular wall; Double-outlet right ventricle; Increase endocardium and myocardium spacing; Impaired septal formation, hypertrabeculation; Increased proliferation	Maintained fetal gene expression (no increase in αMHC, ANF not repressed); Increase in Notch signaling	21402699
	Jmjd2a	Gain and loss of function	Mouse	KO attenuates hypertrophic. Overexpression exacerbates hypertrophy	FHL1	21555854
	Jmjd3	Knockout	mESC	Embryonic lethal by E6.5; Impaired *in vitro* mesoderm and endothelial differentiation in embryonic stem cells	Downregulation of mesoderm and cardiac markers (Brachyury, Mef2c, αMHC, Nppa)	23856522
	Jmjd6	Knockdown	mESC	Decreased in vitro network formation and sprouting in the spheroid assay during the endothelial differentiation of ESCs	Alters Flt1splicing increasing the levels of its soluble form which inhibits angiogenesis.	21300889
	Jmjd8	Knockdown	mESC	Decreased *in vitro* network formation and sprouting in the spheroid assay during the endothelial differentiation of ESCs	Jmjd8 controls glycolysis and respiration of ECs	27199445
	UTX	Knockout	Mouse	Embryonic stem cells cannot differentiate into beating cardiomyocyte *in vitro*	Repression of cardiac markers (ANF, α-cardiac actin, Myh6)	22192413
				Embryonic lethal by E10.5; Impaired looping, cardiomyocyte differentiation	Downregulation of cardiac genes (ANF, Mlc2v)	
Histone Deacetylase	Hdac1, Hdac2	Knockdown	Adult Mouse Ventricular Myocytes	--	Stimulates PAI-1 Expression	28174211
	Hdac6	Knockdown	Rat Cardiomyocytes	--	Upregulation of HDAC10 mRNA expression.	21539845
Protein Kinase	CK2α1	Overexpression	Mouse	Interventricular septum and the left ventricular wall thickness increase	α-actin or β-tubulin upregulation HDAC2 phosphorylation	21576649
Transcriptional Regulation	Meis2	Knockdown	Zebrafish	Reduced heart rate Pericardial edema indicative of cardiac failure	Genes involved in heart tube formation and cardiac looping	22981225
	TFIIB	Knockdown	Rat and Mouse Neonatal Cardiac Myocytes	Ameliorate pressure overload hypertrophy	Genes requiring de novo recruitment of pol II	25398966
	Yy1	Overexpression and Knockdown	Neonatal Mouse Ventricular Myocytes	Up-regulation of all genes present during fetal development and increases cell size	HDAC5	18632988

Table 1-1. Gain/Loss of Function Studies on Chromatin Modifiers in the Cardiovascular System.

Family	Gene	Manipulation	Model	Phenotype	Molecular Targets	Reference (PMID)
ATP-dependent Remodelers	Baf60c	Knockdown	Mouse	Embryonic lethal by E11; Impaired looping, trabeculation, short outflow tract; Hypoplasia	Differential patterns of cardiac gene expression (e.g. Nkx2-5, Nppa, Hand1, Hand2 affected)	15525990, 17210915
	Baf180	Knockout	Mouse	Embryonic lethal by E15.5; Reduced heart rate Impaired septal formation; Hypoplasia; Edema	Downregulation of cell growth and proliferation genes; Upregulation of growth arrest-associated genes	15601824
	Baf250a	Knockdown	P19 Cells	–	Upregulation of cardiac markers (e.g. Nkx2-5, Hand1, Acta2)	22621927, 24335282
		Overexpression		–	Downregulation of cardiac markers (e.g. cTnT, Nkx2-5, Gata4)	
		Knockout	Mouse	Embryonic lethal; Impaired trabeculation of right ventricle, no septum formation, outflow tract; Hypoplasia	Downregulation of cardiac transcription factors (Nkx2-5, Bmp10, Mef2c)	
	Brg1	Cardiac-specific knockout	Zebrafish	Embryonic lethal by E11.5; Impaired looping, no septum formation; Thin ventricular walls due to hypoplasia Embryonic lethal by E11.5; Double-outlet right ventricle; Decreased ventricular chamber size; Impaired looping, septal formation	Downregulation of Bmp1; Increase in αMHC to βMHC ratio Downregulation of cardiac genes (e.g. Nppa, Tbx5, Bmp10)	20596014
		Ventricle-specific knockout	Zebrafish	Neonatal heterozygotes show dilation chambers; impaired septal formation; impaired electrical function	Differential patterns of cardiac gene expression (e.g. Bmp4, Tbx2)	21304516
		Null mutation	Zebrafish	Arrhythmia Hypoplasia	Required for the activation of the Wt1 locus	28737171
		Vascular-specific mutations	Mouse	No evidence of defective angiogenesis. Brg1 and Brm double mutant die with multiorgan hemorrhage including small intestine and heart.	-	25904594
	Brm	Global mutation	Mouse	No evidence of defective angiogenesis. Brg1 and Brm double mutant die with multiorgan hemorrhage including small intestine and heart.	-	
	Chd4	Endothelial-specific knockout	Mouse	Embryonic lethal by E12.5; Impaired vascular integrity, reduced extracellular matrix, hemorrhaging; Hypotrabeculation	Downregulation of extracellular proteins (e.g. Type IV collagen, fibronectin); Differential expression of plasmin regulators to result in activation of matrix metalloproteases	24335282, 24348274
		Knockdown	Mouse	–	Upregulation of cardiac genes (e.g. cTnT, Bmp10, Gata4)	
	Chd7	Knockout	Mouse	Embryonic lethal by E10.5; Interrupted aortic arch	-	17334657, 19855134, 22363697
		Knockdown	Zebrafish	Impaired looping; Pericardial edema	-	
	Dpf3	Knockdown	Zebrafish	Impaired looping; Weak contractility Disrupted myofibril organization	Upregulation of genes involved in metabolic processes. Downregulation of genes involved in ion/electron transport and cell homeostasis	18765789
	Pontin	Knockdown	Zebrafish	Hyperplasia	-	12464178
	Reptin	Active mutant	Zebrafish	Impaired contractility;	-	
				Ventricular hypertrophy due to hyperplasia		
DNA Methylation	Dnmt1	Knockout	Mouse	Embryonic lethal by E11; Underdeveloped heart	-	16066615
	Dnmt3a	Knockout	Mouse cardiomyocytes and hESC	No effect on higher order chromatin	-	29162810
	Dnmt3b	Knockout	Mouse cardiomyocytes and hESC		-	
		Knockout	Mouse	Embryonic lethal by E16.5 Impaired septum formation; Hemorrhaging	-	16501171
		Targeted point mutations	Mouse	Normal development, small body size Thickening of myocardium in some animals	-	
		Cardiac specific knockout	Mouse	No effects in basal conditions. Blunted hypertrophic response after TAC or ISO treatment	Myh7 is alternatively spliced after Dnmt3 depletion	25784084
	MeCP2	Cardiac-specific overexpression	Mouse	Embryonic lethal by E15 Ventricular wall and septal hypertrophy due to hyperplasia	Changes in cardiac genes (e.g. increases in ANF, Mlc2v, Nkx2-5; decreases in βMHC, α-cardiac actin, Tbx5)	20203171
Chromatin Structural Proteins	Ctcf	Cardiac specific knockout	Mouse (Adult)	Hypertrophy and heart failure	Alteration of TAD boundary strength. Chromatin loop disruption. Decrease of the local intra-chromatin interactions. Alteration of enhancer-promoter interactions.	28802249
		Cardiac specific knockout	Mouse (Embryo)	Embryonic lethality after E12.5	Genes involved in mitochondrial function and the lncA cluster.	28846746
		Knockdown	mESC	Cell death after 4 days of depletion	Loss of TAD insulation.	28525758
		Gain and loss of function	C2C12	CTCF induces cell myogenic differentiation	Myogenic development genes (Myf5, MyoD and Myogenin) and Wnt signaling.	21288905
		Gain and loss o function	Zebrafish	CTCF induces myogenesis *in vivo*		
	Hmga2	Overexpression	P19CL6 cells	Efficient differentiation into cardiomyocytes	Upregulation of cardiac markers (Nkx2-5, ANP, Gata4, Mef2c, Mlc2v)	18425117
		Knockdown		Blocked cardiomyocyte differentiation	Downregulation of cardiac markers; Mesodermal markers not affected	
		Dominant negative mutation	Xenopus	Reduced heart size	Downregulation of Nkx2-5	
		Knockdown		–	Downregulation of Nkx2-5	
	Lamin A/C	Knockout	Mouse	Reduced growth rate (postnatal 2wks); Dilated cardiomyopathy and impaired ventricular contractile function (postnatal 4-6wks)	Upregulation of ANP and BNP; Downregulation of SERCA2a; Elongated nuclear shape, disrupted chromatin structure	14755333
	Nucleolin	Knockdown	Zebrafish	Edema; Small ventricular size; Impaired looping Decrease in cardiac output	Upregulation of p53; Decreased rRNA expression; Ventricular bmp4 not repressed; Increase in H3K9me3	24077883
		Overexpression		Edema; Impaired looping, dorsal-ventral axis formation	Upregulation of p53, ventricular Bmp4	
Lymphocyte Antigen	Ly75	Knockdown	Zebrafish	Heart failure with dilation of the atrium and reduced ventricular contractility.	-	23341106
[G protein]-coupled receptor	Adora2a	Knockdown	Zebrafish	Severe heart failure with progressively decreasing ventricular contractility	-	23341106
BET inhibitors	BET inhibitors	JQ1 inhibition	Rat and Mouse Cardiomyocytes	Left ventricular hypertrophy blocking	Calcineurin, NFeB, and GATA4	23939492, 23911322
		JQ1 inhibition	iPSC-CMs	Agonist-induced hypertrophy blocking	NFκB and TGF-β signaling networks	28515341

Table 1-1 (continued). Gain/Loss of Function Studies on Chromatin Modifiers in the Cardiovascular System.

LncRNA	Manipulation	Model	Phenotype	Molecular targets	Reference (PMID)
Braveheart	Knockout	mESC	Necessary for cardiovascular lineage commitment of embryonic stem cells	MesP1 and PRC2	23352431
Chaer	Knockout	Mouse	No noticeable phenotype Cardiac hypertrophy in response to pressure overload was significantly attenuated in the KO	PRC2	27618650
Chast	Loss and gain of function	Mouse	Overexpression induces hypertrophy, increased cardiomyocyte size and fibrosis. Silencing prevents and attenuates TAC-induced pathological cardiac remodeling	Chast activates Bnp, β-Mhc and Ctgf Chast negatively regulated Pleckstrin homology domain-containing protein family M member 1	26888430
Fendrr	Knockout	Mouse	Omphalocele and blood accumulation in the right heart chamber. Embryonic lethal around E13.75	PRC2, E2H2 and SU212	23369715
Hand2	Knockout	Mouse	Right ventricle hypoplasia and embryonic lethality	RNA polymerase II elongation	27783597
Mantis	Knockdown	HDLEC	Regulates endothelial angiogenesis	Brg1, SOX18, SMAD6, and COUP-TFII	28351900
Malat-1	Knockout	Mouse	No effect in cardiac hypertrophy upon pressure overload	mRNA splicing	26919721
Mhrt779	Overexpression	Mouse	Heart protection from hypertrophy and failure	It prevents Brg1 from recognizing its genomic DNA targets	25119045
Uph	Knockdown	Mouse	Right ventricular hypoplasia and embryonic lethality	Hand2	27783597

Table 1-1 (continued). Gain/Loss of Function Studies on Chromatin Modifiers in the Cardiovascular System.

Aco2	Mertk	Acss1	Pxk
Adcy6	Meis2	Kptn	Man1c1
Arvcf	Ndufs4	Med26	Arhgap35
Bad	Nxn	Sik3	Klc3
Bcat2	Serpinf2	Pitpnc1	Pla2g3
Bcl2	Ptp4a3	Dpep3	Fscn2
Bmpr2	Ptpn14	Nkx6-3	Setbp1
Bsn	Pxn	St5	Kcng2
Cct6a	Rai1	Klhl2	Pik3c2b
Cdh16	Rara	Ankrd17	D430041D05Rik
Ift81	Rgl2	Arid1a	2900026A02Rik
Celsr1	Nptn	Cnnm2	Map7d1
Inadl	Tgfbr3	Deptor	Mpg
Cnih2	Tns1	Phlpp1	Gm20735
Col13a1	Cdh23	Ehd4	Ablim3
Col14a1	Vav2	Abtb2	March3
Dapk2	Fmnl3	Slco2b1	Mical2
Des	Car14	Itga9	Lgr6
Dnase1l3	Pappa2	Bre	Zc3h4
Ebf1	Limd1	Ston2	Btbd16
Ehd1	Ldlrad4	Znrf1	Rnf150
Eif4ebp1	Nudt3	Socs7	Lrch1
Gab1	Bcam	Rassf3	6330403A02Rik
Gcgr	Pnkp	Zbtb7c	C730036E19Rik
Gnb1	Slc22a17	Pde2a	Itprip
Grk5	Fam129a	Daam1	Ly6g6f
Grb2	Osbpl1a	Rhot2	Trabd2b
Hhex	Dpysl5	Slc43a2	Gm15441
Itgb5	Nicn1	Stxbp6	Mir1946b
Kcnab3	Map1lc3b	Susd6	Mir1198
Meis1	Wipf2	Mfsd7c	Cbarp

Table 1-2. List of genes with detected significant (q < 0.01 Fit-Hi-C interactions between transcription start and end sites (shown in Figure 1-3).

Chapter 1: Bibliography

1. Kornberg RD. Structure of chromatin. *Annual review of biochemistry*. 1977;46:931-54.

2. Luger K, Mader AW, Richmond RK, Sargent DF and Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*. 1997;389:251-60.

3. Suto RK, Clarkson MJ, Tremethick DJ and Luger K. Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat Struct Biol*. 2000;7:1121-4.

4. Sekulic N, Bassett EA, Rogers DJ and Black BE. The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. *Nature*. 2010;467:347-51.

5. Buschbeck M and Hake SB. Variants of core histones and their roles in cell fate decisions, development and cancer. *Nature reviews Molecular cell biology*. 2017;18:299-314.

6. Talbert PB and Henikoff S. Histone variants on the move: substrates for chromatin dynamics. *Nature reviews Molecular cell biology*. 2017;18:115-126.

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

8. Franklin S, Chen H, Mitchell-Jordan S, Ren S, Wang Y and 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*. 2012;11:M111 014258.

9. Segal E and Widom J. What controls nucleosome positions? *Trends in genetics : TIG*. 2009;25:335-43.

10. Crane-Robinson C. Linker histones: History and current perspectives. *Biochim Biophys Acta*. 2016;1859:431-5.

11. Simpson RT. Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. *Biochemistry*. 1978;17:5524-31.

12. Allan J, Hartman PG, Crane-Robinson C and Aviles FX. The structure of histone H1 and its location in chromatin. *Nature*. 1980;288:675-9.

13. Ramakrishnan V, Finch JT, Graziano V, Lee PL and Sweet RM. Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature*. 1993;362:219-23.

14. Fyodorov DV, Zhou BR, Skoultchi AI and Bai Y. Emerging roles of linker histones in regulating chromatin structure and function. *Nature reviews Molecular cell biology*. 2017.

15. Phillips JE and Corces VG. CTCF: master weaver of the genome. *Cell*. 2009;137:1194-211.

16. Fedele M, Fidanza V, Battista S, Pentimalli F, Klein-Szanto AJ, Visone R, De Martino I, Curcio A, Morisco C, Del Vecchio L, Baldassarre G, Arra C, Viglietto G, Indolfi C, Croce CM and Fusco A. Haploinsufficiency of the Hmga1 gene causes cardiac hypertrophy and myelo-lymphoproliferative disorders in mice. *Cancer Res*. 2006;66:2536-43.

17. Strahl BD and Allis CD. The language of covalent histone modifications. *Nature*. 2000;403:41-5.

18. Turner BM. Histone acetylation and control of gene expression. *J Cell Sci*. 1991;99 (Pt 1):13-20.

19. Stedman E and Stedman E. Cell specificity of histones. *Nature*. 1950;166:780-1.

20. Allfrey VG, Faulkner R and Mirsky AE. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proc Natl Acad Sci U S A*. 1964;51:786-94.

21. Zhao Y and Garcia BA. Comprehensive Catalog of Currently Documented Histone Modifications. *Cold Spring Harb Perspect Biol*. 2015;7:a025064.

22. Zentner GE and Henikoff S. High-resolution digital profiling of the epigenome. *Nat Rev Genet*. 2014;15:814-27.

23. Rhee HS and Pugh BF. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell*. 2011;147:1408-19.

24. Skene PJ and Henikoff S. A simple method for generating high-resolution maps of genome-wide protein binding. *Elife*. 2015;4:e09225.

25. Voigt P, LeRoy G, Drury WJ, 3rd, Zee BM, Son J, Beck DB, Young NL, Garcia BA and Reinberg D. Asymmetrically modified nucleosomes. *Cell*. 2012;151:181-93.

26. Sabari BR, Zhang D, Allis CD and Zhao Y. Metabolic regulation of gene expression through histone acylations. *Nature reviews Molecular cell biology*. 2017;18:90-101.

27. 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 and Haldar SM. BET bromodomains mediate transcriptional pause release in heart failure. *Cell*. 2013;154:569-82.

28. Spiltoir JI, Stratton MS, Cavasin MA, Demos-Davies K, Reid BG, Qi J, Bradner JE and McKinsey TA. BET acetyl-lysine binding proteins control pathological cardiac hypertrophy. *J Mol Cell Cardiol*. 2013;63:175-9.

29. Duan Q, McMahon S, Anand P, Shah H, Thomas S, Salunga HT, Huang Y, Zhang R, Sahadevan A, Lemieux ME, Brown JD, Srivastava D, Bradner JE, McKinsey TA and Haldar SM. BET bromodomain inhibition suppresses innate inflammatory and profibrotic transcriptional networks in heart failure. *Science translational medicine*. 2017;9.

30. Haldar SM and McKinsey TA. BET-ting on chromatin-based therapeutics for heart failure. *J Mol Cell Cardiol*. 2014;74:98-102.

31. Wang Y, Guo YR, Liu K, Yin Z, Liu R, Xia Y, Tan L, Yang P, Lee JH, Li XJ, Hawke D, Zheng Y, Qian X, Lyu J, He J, Xing D, Tao YJ and Lu Z. KAT2A coupled with the alpha-KGDH complex acts as a histone H3 succinyltransferase. *Nature*. 2017;552:273-277.

32. Liew CC and Sole MJ. Nuclear proteins in the heart of the cardiomyopathic Syrian hamster. Fractionation of phenol-soluble nonhistone proteins by two-dimensional polyacrylamide gel electrophoresis. *Circ Res*. 1978;42:628-36.

33. Kislinger T, Cox B, Kannan A, Chung C, Hu P, Ignatchenko A, Scott MS, Gramolini AO, Morris Q, Hallett MT, Rossant J, Hughes TR, Frey B and Emili A. Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. *Cell*. 2006;125:173-86.

34. Frazer KA, Eskin E, Kang HM, Bogue MA, Hinds DA, Beilharz EJ, Gupta RV, Montgomery J, Morenzoni MM, Nilsen GB, Pethiyagoda CL, Stuve LL, Johnson FM, Daly MJ, Wade CM and Cox DR. A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. *Nature*. 2007;448:1050-3.

35. Pickering CM, Grady C, Medvar B, Emamian M, Sandoval PC, Zhao Y, Yang CR, Jung HJ, Chou CL and Knepper MA. Proteomic profiling of nuclear fractions from native renal inner medullary collecting duct cells. *Physiological genomics*. 2016;48:154-66.

36. Zheng Y, Huang X and Kelleher NL. Epiproteomics: quantitative analysis of histone marks and codes by mass spectrometry. *Curr Opin Chem Biol*. 2016;33:142-50.

37. Dann GP, Liszczak GP, Bagert JD, Muller MM, Nguyen UTT, Wojcik F, Brown ZZ, Bos J, Panchenko T, Pihl R, Pollock SB, Diehl KL, Allis CD and Muir TW. ISWI chromatin remodellers sense nucleosome modifications to determine substrate preference. *Nature*. 2017;548:607-611.

38. Nguyen UT, Bittova L, Muller MM, Fierz B, David Y, Houck-Loomis B, Feng V, Dann GP and Muir TW. Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries. *Nat Methods*. 2014;11:834-40.

39. Okamoto H, Fujioka Y, Takahashi A, Takahashi T, Taniguchi T, Ishikawa Y and Yokoyama M. Trichostatin A, an inhibitor of histone deacetylase, inhibits smooth muscle cell proliferation via induction of p21(WAF1). *J Atheroscler Thromb*. 2006;13:183-91.

40. Kong X, Fang M, Li P, Fang F and Xu Y. HDAC2 deacetylates class II transactivator and suppresses its activity in macrophages and smooth muscle cells. *J Mol Cell Cardiol*. 2009;46:292-9.

41. Zheng XX, Zhou T, Wang XA, Tong XH and Ding JW. Histone deacetylases and atherosclerosis. *Atherosclerosis*. 2015;240:355-66.

42. Cardinale JP, Sriramula S, Pariaut R, Guggilam A, Mariappan N, Elks CM and Francis J. HDAC inhibition attenuates inflammatory, hypertrophic, and hypertensive responses in spontaneously hypertensive rats. *Hypertension*. 2010;56:437-44.

43. Lee TM, Lin MS and Chang NC. Inhibition of histone deacetylase on ventricular remodeling in infarcted rats. *Am J Physiol Heart Circ Physiol*. 2007;293:H968-77.

44. Granger A, Abdullah I, Huebner F, Stout A, Wang T, Huebner T, Epstein JA and Gruber PJ. Histone deacetylase inhibition reduces myocardial ischemia-reperfusion injury in mice. *FASEB J*. 2008;22:3549-60.

45. Zhao TC, Cheng G, Zhang LX, Tseng YT and Padbury JF. Inhibition of histone deacetylases triggers pharmacologic preconditioning effects against myocardial ischemic injury. *Cardiovasc Res*. 2007;76:473-81.

46. Aune SE, Herr DJ, Mani SK and Menick DR. Selective inhibition of class I but not class IIb histone deacetylases exerts cardiac protection from ischemia reperfusion. *J Mol Cell Cardiol*. 2014;72:138-45.

47. McKinsey TA and Olson EN. Dual roles of histone deacetylases in the control of cardiac growth. *Novartis Found Symp*. 2004;259:132-41; discussion 141-5, 163-9.

48. Cao DJ, Wang ZV, Battiprolu PK, Jiang N, Morales CR, Kong Y, Rothermel BA, Gillette TG and Hill JA. Histone deacetylase (HDAC) inhibitors attenuate cardiac hypertrophy by suppressing autophagy. *Proc Natl Acad Sci U S A*. 2011;108:4123-8.

49. Kook H, Lepore JJ, Gitler AD, Lu MM, Wing-Man Yung W, Mackay J, Zhou R, Ferrari V, Gruber P and Epstein JA. Cardiac hypertrophy and histone deacetylase-dependent transcriptional repression mediated by the atypical homeodomain protein Hop. *J Clin Invest*. 2003;112:863-71.

50. Zhang CL, McKinsey TA, Chang S, Antos CL, Hill JA and Olson EN. Class II histone deacetylases act as signal-responsive repressors of cardiac hypertrophy. *Cell*. 2002;110:479-88.

51. Backs J, Backs T, Bezprozvannaya S, McKinsey TA and Olson EN. Histone deacetylase 5 acquires calcium/calmodulin-dependent kinase II responsiveness by oligomerization with histone deacetylase 4. *Mol Cell Biol*. 2008;28:3437-45.

52. Zhang T, Kohlhaas M, Backs J, Mishra S, Phillips W, Dybkova N, Chang S, Ling H, Bers DM, Maier LS, Olson EN and Brown JH. CaMKIIdelta isoforms differentially affect calcium handling but similarly regulate HDAC/MEF2 transcriptional responses. *J Biol Chem*. 2007;282:35078-87.

53. Trivedi CM, Luo Y, Yin Z, Zhang M, Zhu W, Wang T, Floss T, Goettlicher M, Noppinger PR, Wurst W, Ferrari VA, Abrams CS, Gruber PJ and Epstein JA. Hdac2 regulates the cardiac hypertrophic response by modulating Gsk3 beta activity. *Nature medicine*. 2007;13:324-31.

54. McKinsey TA. Isoform-selective HDAC inhibitors: closing in on translational medicine for the heart. *J Mol Cell Cardiol*. 2011;51:491-6.

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

56. Matsushima S and Sadoshima J. The role of sirtuins in cardiac disease. *Am J Physiol Heart Circ Physiol*. 2015;309:H1375-89.

57. Yiew KH, Chatterjee TK, Hui DY and Weintraub NL. Histone Deacetylases and Cardiometabolic Diseases. *Arterioscler Thromb Vasc Biol*. 2015;35:1914-9.

58. Byrne MM, Murphy RT and Ryan AW. Epigenetic modulation in the treatment of atherosclerotic disease. *Frontiers in genetics*. 2014;5:364.

59. Voelter-Mahlknecht S. Epigenetic associations in relation to cardiovascular prevention and therapeutics. *Clin Epigenetics*. 2016;8:4.

60. Wang Y, Miao X, Liu Y, Li F, Liu Q, Sun J and Cai L. Dysregulation of histone acetyltransferases and deacetylases in cardiovascular diseases. *Oxid Med Cell Longev*. 2014;2014:641979.

61. Pons D, de Vries FR, van den Elsen PJ, Heijmans BT, Quax PH and Jukema JW. Epigenetic histone acetylation modifiers in vascular remodelling: new targets for therapy in cardiovascular disease. *European heart journal*. 2009;30:266-77.

62. Zhang Y, Qiu J, Wang X, Zhang Y and Xia M. AMP-activated protein kinase suppresses endothelial cell inflammation through phosphorylation of transcriptional coactivator p300. *Arterioscler Thromb Vasc Biol*. 2011;31:2897-908.

63. Morimoto T, Sunagawa Y, Kawamura T, Takaya T, Wada H, Nagasawa A, Komeda M, Fujita M, Shimatsu A, Kita T and Hasegawa K. The dietary compound curcumin inhibits p300 histone acetyltransferase activity and prevents heart failure in rats. *J Clin Invest*. 2008;118:868-78.

64. Gusterson RJ, Jazrawi E, Adcock IM and Latchman DS. The transcriptional co-activators CREB-binding protein (CBP) and p300 play a critical role in cardiac hypertrophy that is dependent on their histone acetyltransferase activity. *J Biol Chem*. 2003;278:6838-47.

65. Qiao W, Zhang W, Gai Y, Zhao L and Fan J. The histone acetyltransferase MOF overexpression blunts cardiac hypertrophy by targeting ROS in mice. *Biochem Biophys Res Commun*. 2014;448:379-84.

66. Yi X, Tao Y, Lin X, Dai Y, Yang T, Yue X, Jiang X, Li X, Jiang DS, Andrade KC and Chang J. Histone methyltransferase Setd2 is critical for the proliferation and differentiation of myoblasts. *Biochim Biophys Acta*. 2017;1864:697-707.

67. Franklin S, Kimball T, Rasmussen TL, Rosa-Garrido M, Chen H, Tran T, Miller MR, Gray R, Jiang S, Ren S, Wang Y, Tucker HO and Vondriska TM. The chromatin-binding protein Smyd1 restricts adult mammalian heart growth. *Am J Physiol Heart Circ Physiol*. 2016;311:H1234-H1247.

68. Franklin S, Kimball T, Rasmussen TL, Rosa Garrido M, Chen H, Tran T, Miller MR, Gray R, Jiang S, Ren S, Wang Y, Tucker HO and Vondriska TM. The chromatin binding protein Smyd1 restricts adult mammalian heart growth. *Am J Physiol Heart Circ Physiol*. 2016;In Press.

69. Tan X, Rotllant J, Li H, De Deyne P and Du SJ. SmyD1, a histone methyltransferase, is required for myofibril organization and muscle contraction in zebrafish embryos. *Proc Natl Acad Sci U S A*. 2006;103:2713-8.

70. Voelkel T, Andresen C, Unger A, Just S, Rottbauer W and Linke WA. Lysine methyltransferase Smyd2 regulates Hsp90-mediated protection of the sarcomeric titin springs and cardiac function. *Biochim Biophys Acta*. 2013;1833:812-22.

71. Just S, Meder B, Berger IM, Etard C, Trano N, Patzel E, Hassel D, Marquart S, Dahme T, Vogel B, Fishman MC, Katus HA, Strahle U and Rottbauer W. The myosin-interacting protein SMYD1 is essential for sarcomere organization. *J Cell Sci*. 2011;124:3127-36.

72. Rasmussen TL, Ma Y, Park CY, Harriss J, Pierce SA, Dekker JD, Valenzuela N, Srivastava D, Schwartz RJ, Stewart MD and Tucker HO. Smyd1 facilitates heart development by antagonizing oxidative and ER stress responses. *PLoS One*. 2015;10:e0121765.

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

74. Papait R, Serio S, Pagiatakis C, Rusconi F, Carullo P, Mazzola M, Salvarani N, Miragoli M and Condorelli G. Histone Methyltransferase G9a Is Required for Cardiomyocyte Homeostasis and Hypertrophy. *Circulation*. 2017;136:1233-1246.

75. Okabe J, Orlowski C, Balcerczyk A, Tikellis C, Thomas MC, Cooper ME and El-Osta A. Distinguishing hyperglycemic changes by Set7 in vascular endothelial cells. *Circ Res*. 2012;110:1067-76.

76. Keating ST and El-Osta A. Epigenetics and metabolism. *Circ Res*. 2015;116:715-36.

77. Zhang QJ, Chen HZ, Wang L, Liu DP, Hill JA and Liu ZP. The histone trimethyllysine demethylase JMJD2A promotes cardiac hypertrophy in response to hypertrophic stimuli in mice. *J Clin Invest*. 2011;121:2447-56.

78. Ono T, Kamimura N, Matsuhashi T, Nagai T, Nishiyama T, Endo J, Hishiki T, Nakanishi T, Shimizu N, Tanaka H, Ohta S, Suematsu M, Ieda M, Sano M, Fukuda K and Kaneda R. The histone 3 lysine 9 methyltransferase inhibitor chaetocin improves prognosis in a rat model of high salt diet-induced heart failure. *Sci Rep*. 2017;7:39752.

79. Ai S, Yu X, Li Y, Peng Y, Li C, Yue Y, Tao G, Li C, Pu WT and He A. Divergent Requirements for EZH1 in Heart Development Versus Regeneration. *Circ Res*. 2017;121:106-112.

80. Delgado-Olguin P, Dang LT, He D, Thomas S, Chi L, Sukonnik T, Khyzha N, Dobenecker MW, Fish JE and Bruneau BG. Ezh2-mediated repression of a transcriptional pathway upstream of Mmp9 maintains integrity of the developing vasculature. *Development*. 2014;141:4610-7.

81. Mitic T, Caporali A, Floris I, Meloni M, Marchetti M, Urrutia R, Angelini GD and Emanueli C. EZH2 modulates angiogenesis in vitro and in a mouse model of limb ischemia. *Mol Ther*. 2015;23:32-42.

82. Kim E, Bisson WH, Lohr CV, Williams DE, Ho E, Dashwood RH and Rajendran P. Histone and Non-Histone Targets of Dietary Deacetylase Inhibitors. *Curr Top Med Chem*. 2016;16:714-31.

83. Nian H, Delage B, Ho E and Dashwood RH. Modulation of histone deacetylase activity by dietary isothiocyanates and allyl sulfides: studies with sulforaphane and garlic organosulfur compounds. *Environ Mol Mutagen*. 2009;50:213-21.

84. Jeong MY, Lin YH, Wennersten SA, Demos-Davies KM, Cavasin MA, Mahaffey JH, Monzani V, Saripalli C, Mascagni P, Reece TB, Ambardekar AV, Granzier HL, Dinarello CA and McKinsey TA. Histone deacetylase activity governs diastolic dysfunction through a nongenomic mechanism. *Science translational medicine*. 2018;10.

85. Ho L and Crabtree GR. Chromatin remodelling during development. *Nature*. 2010;463:474-84.

86. Narlikar GJ, Sundaramoorthy R and Owen-Hughes T. Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell*. 2013;154:490-503.

87. Clapier CR, Iwasa J, Cairns BR and Peterson CL. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nature reviews Molecular cell biology*. 2017;18:407-422.

88. Harikumar A and Meshorer E. Chromatin remodeling and bivalent histone modifications in embryonic stem cells. *EMBO Rep*. 2015;16:1609-19.

89. Wiley MM, Muthukumar V, Griffin TM and Griffin CT. SWI/SNF chromatin-remodeling enzymes Brahma-related gene 1 (BRG1) and Brahma (BRM) are dispensable in multiple models of postnatal angiogenesis but are required for vascular integrity in infant mice. *J Am Heart Assoc*. 2015;4.

90. Willis MS, Homeister JW, Rosson GB, Annayev Y, Holley D, Holly SP, Madden VJ, Godfrey V, Parise LV and Bultman SJ. Functional redundancy of SWI/SNF catalytic subunits in maintaining vascular endothelial cells in the adult heart. *Circ Res*. 2012;111:e111-22.

91. Xiao C, Gao L, Hou Y, Xu C, Chang N, Wang F, Hu K, He A, Luo Y, Wang J, Peng J, Tang F, Zhu X and Xiong JW. Chromatin-remodelling factor Brg1 regulates myocardial proliferation and regeneration in zebrafish. *Nature communications*. 2016;7:13787.

92. Alexander JM, Hota SK, He D, Thomas S, Ho L, Pennacchio LA and Bruneau BG. Brg1 modulates enhancer activation in mesoderm lineage commitment. *Development*. 2015;142:1418-30.

93. Lickert H, Takeuchi JK, Von Both I, Walls JR, McAuliffe F, Adamson SL, Henkelman RM, Wrana JL, Rossant J and Bruneau BG. Baf60c is essential for function of BAF chromatin remodelling complexes in heart development. *Nature*. 2004;432:107-12.

94. Hang CT, Yang J, Han P, Cheng HL, Shang C, Ashley E, Zhou B and Chang CP. Chromatin regulation by Brg1 underlies heart muscle development and disease. *Nature*. 2010;466:62-7.

95. 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 and Bruneau BG. Chromatin remodelling complex dosage modulates transcription factor function in heart development. *Nature communications*. 2011;2:187.

96. Griffin CT, Brennan J and Magnuson T. The chromatin-remodeling enzyme BRG1 plays an essential role in primitive erythropoiesis and vascular development. *Development*. 2008;135:493-500.

97. Zhang M, Chen M, Kim JR, Zhou J, Jones RE, Tune JD, Kassab GS, Metzger D, Ahlfeld S, Conway SJ and Herring BP. SWI/SNF complexes containing Brahma or Brahma-related gene 1 play distinct roles in smooth muscle development. *Mol Cell Biol*. 2011;31:2618-31.

98. Fang F, Chen D, Yu L, Dai X, Yang Y, Tian W, Cheng X, Xu H, Weng X, Fang M, Zhou J, Gao Y, Chen Q and Xu Y. Proinflammatory stimuli engage Brahma related gene 1 and Brahma in endothelial injury. *Circ Res*. 2013;113:986-96.

99. Guo JU, Su Y, Shin JH, Shin J, Li H, Xie B, Zhong C, Hu S, Le T, Fan G, Zhu H, Chang Q, Gao Y, Ming GL and Song H. Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nature neuroscience*. 2014;17:215-22.

100. Zhang D, Wu B, Wang P, Wang Y, Lu P, Nechiporuk T, Floss T, Greally JM, Zheng D and Zhou B. Non-CpG methylation by DNMT3B facilitates REST binding and gene silencing in developing mouse hearts. *Nucleic Acids Res*. 2017;45:3102-3115.

101. Nothjunge S, Nuhrenberg TG, Gruning BA, Doppler SA, Preissl S, Schwaderer M, Rommel C, Krane M, Hein L and Gilsbach R. DNA methylation signatures follow preformed chromatin compartments in cardiac myocytes. *Nature communications*. 2017;8:1667.

102. Orozco LD, Morselli M, Rubbi L, Guo W, Go J, Shi H, Lopez D, Furlotte NA, Bennett BJ, Farber CR, Ghazalpour A, Zhang MQ, Bahous R, Rozen R, Lusis AJ and Pellegrini M. Epigenome-wide association of liver methylation patterns and complex metabolic traits in mice. *Cell metabolism*. 2015;21:905-17.

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

104. Teschendorff AE and Relton CL. Statistical and integrative system-level analysis of DNA methylation data. *Nat Rev Genet*. 2018;19:129-147.

105. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13:484-92.

106. Aavik E, Lumivuori H, Leppanen O, Wirth T, Hakkinen SK, Brasen JH, Beschorner U, Zeller T, Braspenning M, van Criekinge W, Makinen K and Yla-Herttuala S. Global DNA methylation analysis of human atherosclerotic plaques reveals extensive genomic hypomethylation and reactivation at imprinted locus 14q32 involving induction of a miRNA cluster. *European heart journal*. 2015;36:993-1000.

107. Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA, Bibikova M and Esteller M. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics : official journal of the DNA Methylation Society*. 2011;6:692-702.

108. Richard MA, Huan T, Ligthart S, Gondalia R, Jhun MA, Brody JA, Irvin MR, Marioni R, Shen J, Tsai PC, Montasser ME, Jia Y, Syme C, Salfati EL, Boerwinkle E, Guan W, Mosley TH, Jr., Bressler J, Morrison AC, Liu C, Mendelson MM, Uitterlinden AG, van Meurs JB, Consortium B, Franco OH, Zhang G, Li Y, Stewart JD, Bis JC, Psaty BM, Chen YI, Kardia SLR, Zhao W, Turner ST, Absher D, Aslibekyan S, Starr JM, McRae AF, Hou L, Just AC, Schwartz JD, Vokonas PS, Menni C, Spector TD, Shuldiner A, Damcott CM, Rotter JI, Palmas W, Liu Y, Paus T, Horvath S, O'Connell JR, Guo X, Pausova Z, Assimes TL, Sotoodehnia N, Smith JA, Arnett DK, Deary IJ, Baccarelli AA, Bell JT, Whitsel E, Dehghan A, Levy D and Fornage M. DNA Methylation Analysis Identifies Loci for Blood Pressure Regulation. *American journal of human genetics*. 2017;101:888-902.

109. Mendelson MM, Marioni RE, Joehanes R, Liu C, Hedman AK, Aslibekyan S, Demerath EW, Guan W, Zhi D, Yao C, Huan T, Willinger C, Chen B, Courchesne P, Multhaup M, Irvin MR, Cohain A, Schadt EE, Grove ML, Bressler J, North K, Sundstrom J, Gustafsson S, Shah S, McRae AF, Harris SE, Gibson J, Redmond P, Corley J, Murphy L, Starr JM, Kleinbrink E, Lipovich L, Visscher PM, Wray NR, Krauss RM, Fallin D, Feinberg A, Absher DM, Fornage M, Pankow JS, Lind L, Fox C, Ingelsson E, Arnett DK, Boerwinkle E, Liang L, Levy D and Deary IJ. Association of Body Mass Index with DNA Methylation and Gene Expression in Blood Cells and Relations to Cardiometabolic Disease: A Mendelian Randomization Approach. *PLoS Med*. 2017;14:e1002215.

110. Aslibekyan S, Demerath EW, Mendelson M, Zhi D, Guan W, Liang L, Sha J, Pankow JS, Liu C, Irvin MR, Fornage M, Hidalgo B, Lin LA, Thibeault KS, Bressler J, Tsai MY, Grove ML, Hopkins PN, Boerwinkle E, Borecki IB, Ordovas JM, Levy D, Tiwari HK, Absher DM and Arnett DK. Epigenome-wide study identifies novel methylation loci associated with body mass index and waist circumference. *Obesity (Silver Spring)*. 2015;23:1493-501.

111. Lin H, Yin X, Xie Z, Lunetta KL, Lubitz SA, Larson MG, Ko D, Magnani JW, Mendelson MM, Liu C, McManus DD, Levy D, Ellinor PT and Benjamin EJ. Methylome-wide Association Study of Atrial Fibrillation in Framingham Heart Study. *Sci Rep*. 2017;7:40377.

112. Ligthart S, Marzi C, Aslibekyan S, Mendelson MM, Conneely KN, Tanaka T, Colicino E, Waite LL, Joehanes R, Guan W, Brody JA, Elks C, Marioni R, Jhun MA, Agha G, Bressler J, Ward-Caviness CK, Chen BH, Huan T, Bakulski K, Salfati EL, Investigators W-E, Fiorito G, Disease CeoCH, Wahl S, Schramm K, Sha J, Hernandez DG, Just AC, Smith JA, Sotoodehnia N, Pilling LC, Pankow JS, Tsao PS, Liu C, Zhao W, Guarrera S, Michopoulos VJ, Smith AK, Peters MJ, Melzer D, Vokonas P, Fornage M, Prokisch H, Bis JC, Chu AY, Herder C, Grallert H, Yao C, Shah S, McRae AF, Lin H, Horvath S, Fallin D, Hofman A, Wareham NJ, Wiggins KL, Feinberg AP, Starr JM, Visscher PM, Murabito JM, Kardia SL, Absher DM, Binder EB, Singleton AB, Bandinelli S, Peters A, Waldenberger M, Matullo G, Schwartz JD, Demerath EW, Uitterlinden AG, van Meurs JB, Franco OH, Chen YI, Levy D, Turner ST, Deary IJ, Ressler KJ, Dupuis J, Ferrucci L, Ong KK, Assimes TL, Boerwinkle E, Koenig W, Arnett DK, Baccarelli AA, Benjamin EJ and Dehghan A.

DNA methylation signatures of chronic low-grade inflammation are associated with complex diseases. *Genome Biol*. 2016;17:255.

113. Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai PC, Roetker NS, Just AC, Demerath EW, Guan W, Bressler J, Fornage M, Studenski S, Vandiver AR, Moore AZ, Tanaka T, Kiel DP, Liang L, Vokonas P, Schwartz J, Lunetta KL, Murabito JM, Bandinelli S, Hernandez DG, Melzer D, Nalls M, Pilling LC, Price TR, Singleton AB, Gieger C, Holle R, Kretschmer A, Kronenberg F, Kunze S, Linseisen J, Meisinger C, Rathmann W, Waldenberger M, Visscher PM, Shah S, Wray NR, McRae AF, Franco OH, Hofman A, Uitterlinden AG, Absher D, Assimes T, Levine ME, Lu AT, Tsao PS, Hou L, Manson JE, Carty CL, LaCroix AZ, Reiner AP, Spector TD, Feinberg AP, Levy D, Baccarelli A, van Meurs J, Bell JT, Peters A, Deary IJ, Pankow JS, Ferrucci L and Horvath S. DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging (Albany NY)*. 2016;8:1844-1865.

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

115. Greco CM, Kunderfranco P, Rubino M, Larcher V, Carullo P, Anselmo A, Kurz K, Carell T, Angius A, Latronico MV, Papait R and Condorelli G. DNA hydroxymethylation controls cardiomyocyte gene expression in development and hypertrophy. *Nature communications*. 2016;7:12418.

116. Movassagh M, Choy MK, Knowles DA, Cordeddu L, Haider S, Down T, Siggens L, Vujic A, Simeoni I, Penkett C, Goddard M, Lio P, Bennett MR and Foo RS. Distinct epigenomic features in end-stage failing human hearts. *Circulation*. 2011;124:2411-22.

117. Movassagh M, Vujic A and Foo R. Genome-wide DNA methylation in human heart failure. *Epigenomics*. 2011;3:103-9.

118. Shen Y, Matsuno Y, Fouse SD, Rao N, Root S, Xu R, Pellegrini M, Riggs AD and Fan G. X-inactivation in female human embryonic stem cells is in a nonrandom pattern and prone to epigenetic alterations. *Proc Natl Acad Sci U S A*. 2008;105:4709-14.

119. Lee JT. Epigenetic regulation by long noncoding RNAs. *Science*. 2012;338:1435-9.

120. Mayer SC, Gilsbach R, Preissl S, Monroy Ordonez EB, Schnick T, Beetz N, Lother A, Rommel C, Ihle H, Bugger H, Ruhle F, Schrepper A, Schwarzer M, Heilmann C, Bonisch U, Gupta SK, Wilpert J, Kretz O, von Elverfeldt D, Orth J, Aktories K, Beyersdorf F, Bode C, Stiller B, Kruger M, Thum T, Doenst T, Stoll M and Hein L. Adrenergic Repression of the Epigenetic Reader MeCP2 Facilitates Cardiac Adaptation in Chronic Heart Failure. *Circ Res*. 2015;117:622-33.

121. Gilsbach R, Schwaderer M, Preissl S, Gruning BA, Kranzhofer D, Schneider P, Nuhrenberg TG, Mulero-Navarro S, Weichenhan D, Braun C, Dressen M, Jacobs AR, Lahm H, Doenst T, Backofen R, Krane M, Gelb BD and Hein L. Distinct epigenetic programs regulate cardiac myocyte development and disease in the human heart in vivo. *Nature communications*. 2018;9:391.

122. Jones PA, Issa JP and Baylin S. Targeting the cancer epigenome for therapy. *Nat Rev Genet*. 2016;17:630-41.

123. Zaidi S, Choi M, Wakimoto H, Ma L, Jiang J, Overton JD, Romano-Adesman A, Bjornson RD, Breitbart RE, Brown KK, Carriero NJ, Cheung YH, Deanfield J, DePalma S, Fakhro KA, Glessner J, Hakonarson H, Italia MJ, Kaltman JR, Kaski J, Kim R, Kline JK, Lee T, Leipzig J, Lopez A, Mane SM, Mitchell LE, Newburger JW, Parfenov M, Pe'er I, Porter G, Roberts AE, Sachidanandam R, Sanders SJ, Seiden HS, State MW, Subramanian S, Tikhonova IR, Wang W, Warburton D, White PS, Williams IA, Zhao H, Seidman JG, Brueckner M, Chung WK, Gelb BD, Goldmuntz E, Seidman CE and Lifton RP. De novo mutations in histone-modifying genes in congenital heart disease. *Nature*. 2013;498:220-3.

124. Bar C, Chatterjee S and Thum T. Long Noncoding RNAs in Cardiovascular Pathology, Diagnosis, and Therapy. *Circulation*. 2016;134:1484-1499.

125. Poller W, Dimmeler S, Heymans S, Zeller T, Haas J, Karakas M, Leistner DM, Jakob P, Nakagawa S, Blankenberg S, Engelhardt S, Thum T, Weber C, Meder B, Hajjar R and Landmesser U. Non-coding RNAs in cardiovascular diseases: diagnostic and therapeutic perspectives. *European heart journal*. 2017.

126. Uchida S and Dimmeler S. Long noncoding RNAs in cardiovascular diseases. *Circ Res*. 2015;116:737-50.

127. Mathiyalagan P, Keating ST, Du XJ and El-Osta A. Interplay of chromatin modifications and non-coding RNAs in the heart. *Epigenetics : official journal of the DNA Methylation Society*. 2014;9:101-12.

128. Magny EG, Pueyo JI, Pearl FM, Cespedes MA, Niven JE, Bishop SA and Couso JP. Conserved regulation of cardiac calcium uptake by peptides encoded in small open reading frames. *Science*. 2013;341:1116-20.

129. Anderson DM, Anderson KM, Chang CL, Makarewich CA, Nelson BR, McAnally JR, Kasaragod P, Shelton JM, Liou J, Bassel-Duby R and Olson EN. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell*. 2015;160:595-606.

130. Engreitz JM, Pandya-Jones A, McDonel P, Shishkin A, Sirokman K, Surka C, Kadri S, Xing J, Goren A, Lander ES, Plath K and Guttman M. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science*. 2013;341:1237973.

131. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E and Chang HY. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*. 2007;129:1311-23.

132. Zhao J, Ohsumi TK, Kung JT, Ogawa Y, Grau DJ, Sarma K, Song JJ, Kingston RE, Borowsky M and Lee JT. Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol Cell*. 2010;40:939-53.

133. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, Regev A, Lander ES and Rinn JL. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A*. 2009;106:11667-72.

134. Tu S, Yuan GC and Shao Z. The PRC2-binding long non-coding RNAs in human and mouse genomes are associated with predictive sequence features. *Sci Rep*. 2017;7:41669.

135. Beltran M, Yates CM, Skalska L, Dawson M, Reis FP, Viiri K, Fisher CL, Sibley CR, Foster BM, Bartke T, Ule J and Jenner RG. The interaction of PRC2 with RNA or chromatin is mutually antagonistic. *Genome Res*. 2016;26:896-907.

136. Joung J, Engreitz JM, Konermann S, Abudayyeh OO, Verdine VK, Aguet F, Gootenberg JS, Sanjana NE, Wright JB, Fulco CP, Tseng YY, Yoon CH, Boehm JS, Lander ES and Zhang F. Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. *Nature*. 2017;548:343-346.

137. Bohmdorfer G and Wierzbicki AT. Control of Chromatin Structure by Long Noncoding RNA. *Trends in cell biology*. 2015;25:623-632.

138. Grote P, Wittler L, Hendrix D, Koch F, Wahrisch S, Beisaw A, Macura K, Blass G, Kellis M, Werber M and Herrmann BG. The Tissue-Specific lncRNA Fendrr Is an Essential Regulator of Heart and Body Wall Development in the Mouse. *Developmental cell*. 2013;24:206-14.

139. 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 and Boyer LA. Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell*. 2013;152:570-83.

140. Anderson KM, Anderson DM, McAnally JR, Shelton JM, Bassel-Duby R and Olson EN. Transcription of the non-coding RNA upperhand controls Hand2 expression and heart development. *Nature*. 2016;539:433-436.

141. Wang Z, Zhang XJ, Ji YX, Zhang P, Deng KQ, Gong J, Ren S, Wang X, Chen I, Wang H, Gao C, Yokota T, Ang YS, Li S, Cass A, Vondriska TM, Li G, Deb A, Srivastava

D, Yang HT, Xiao X, Li H and Wang Y. A long non-coding RNA defines an epigenetic checkpoint for cardiac hypertrophy *Nature medicine*. 2016.

142. Mathiyalagan P, Okabe J, Chang L, Su Y, Du XJ and El-Osta A. The primary microRNA-208b interacts with Polycomb-group protein, Ezh2, to regulate gene expression in the heart. *Nucleic Acids Res*. 2014;42:790-803.

143. Leisegang MS, Fork C, Josipovic I, Richter FM, Preussner J, Hu J, Miller MJ, Epah J, Hofmann P, Gunther S, Moll F, Valasarajan C, Heidler J, Ponomareva Y, Freiman TM, Maegdefessel L, Plate KH, Mittelbronn M, Uchida S, Kunne C, Stellos K, Schermuly RT, Weissmann N, Devraj K, Wittig I, Boon RA, Dimmeler S, Pullamsetti SS, Looso M, Miller FJ, Jr. and Brandes RP. Long Noncoding RNA MANTIS Facilitates Endothelial Angiogenic Function. *Circulation*. 2017;136:65-79.

144. 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 and Chang CP. A long noncoding RNA protects the heart from pathological hypertrophy. *Nature*. 2014;514:102-6.

145. Michalik KM, You X, Manavski Y, Doddaballapur A, Zornig M, Braun T, John D, Ponomareva Y, Chen W, Uchida S, Boon RA and Dimmeler S. Long noncoding RNA MALAT1 regulates endothelial cell function and vessel growth. *Circ Res*. 2014;114:1389-97.

146. Viereck J, Kumarswamy R, Foinquinos A, Xiao K, Avramopoulos P, Kunz M, Dittrich M, Maetzig T, Zimmer K, Remke J, Just A, Fendrich J, Scherf K, Bolesani E, Schambach A, Weidemann F, Zweigerdt R, de Windt LJ, Engelhardt S, Dandekar T,

Batkai S and Thum T. Long noncoding RNA Chast promotes cardiac remodeling. *Science translational medicine*. 2016;8:326ra22.

147. Sallam T, Jones MC, Gilliland T, Zhang L, Wu X, Eskin A, Sandhu J, Casero D, Vallim TQ, Hong C, Katz M, Lee R, Whitelegge J and Tontonoz P. Feedback modulation of cholesterol metabolism by the lipid-responsive non-coding RNA LeXis. *Nature*. 2016;534:124-8.

148. Sallam T, Jones M, Thomas BJ, Wu X, Gilliland T, Qian K, Eskin A, Casero D, Zhang Z, Sandhu J, Salisbury D, Rajbhandari P, Civelek M, Hong C, Ito A, Liu X, Daniel B, Lusis AJ, Whitelegge J, Nagy L, Castrillo A, Smale S and Tontonoz P. Transcriptional regulation of macrophage cholesterol efflux and atherogenesis by a long noncoding RNA. *Nature medicine*. 2018.

149. Matkovich SJ, Edwards JR, Grossenheider TC, de Guzman Strong C and Dorn GW, 2nd. Epigenetic coordination of embryonic heart transcription by dynamically regulated long noncoding RNAs. *Proc Natl Acad Sci U S A*. 2014.

150. Rajabi M, Kassiotis C, Razeghi P and Taegtmeyer H. Return to the fetal gene program protects the stressed heart: a strong hypothesis. *Heart Fail Rev*. 2007;12:331-43.

151. Ounzain S, Micheletti R, Beckmann T, Schroen B, Alexanian M, Pezzuto I, Crippa S, Nemir M, Sarre A, Johnson R, Dauvillier J, Burdet F, Ibberson M, Guigo R, Xenarios I, Heymans S and Pedrazzini T. Genome-wide profiling of the cardiac transcriptome after myocardial infarction identifies novel heart-specific long non-coding RNAs. *European heart journal*. 2015;36:353-68a.

152. He C, Hu H, Wilson KD, Wu H, Feng J, Xia S, Churko J, Qu K, Chang HY and Wu JC. Systematic Characterization of Long Noncoding RNAs Reveals the Contrasting Coordination of Cis- and Trans-Molecular Regulation in Human Fetal and Adult Hearts. *Circulation Cardiovascular genetics*. 2016;9:110-8.

153. Buske FA, Mattick JS and Bailey TL. Potential in vivo roles of nucleic acid triple-helices. *RNA Biol*. 2011;8:427-39.

154. Felsenfeld G, Davies DR and Rich A. Formation of a three stranded polynucleotide molecule. *J Am Chem Soc*. 1957;79:2023-2024.

155. Shlyueva D, Stampfel G and Stark A. Transcriptional enhancers: from properties to genome-wide predictions. *Nat Rev Genet*. 2014;15:272-86.

156. Calo E and Wysocka J. Modification of enhancer chromatin: what, how, and why? *Mol Cell*. 2013;49:825-37.

157. Heinz S, Romanoski CE, Benner C and Glass CK. The selection and function of cell type-specific enhancers. *Nature reviews Molecular cell biology*. 2015;16:144-54.

158. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, Hoke HA and Young RA. Super-enhancers in the control of cell identity and disease. *Cell*. 2013;155:934-47.

159. Hnisz D, Shrinivas K, Young RA, Chakraborty AK and Sharp PA. A Phase Separation Model for Transcriptional Control. *Cell*. 2017;169:13-23.

160. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G and Greenberg ME. Widespread transcription at neuronal activity-regulated enhancers. *Nature*. 2010;465:182-7.

161. Danko CG, Hyland SL, Core LJ, Martins AL, Waters CT, Lee HW, Cheung VG, Kraus WL, Lis JT and Siepel A. Identification of active transcriptional regulatory elements from GRO-seq data. *Nat Methods*. 2015;12:433-8.

162. Yang XH, Nadadur RD, Hilvering CR, Bianchi V, Werner M, Mazurek SR, Gadek M, Shen KM, Goldman JA, Tyan L, Bekeny J, Hall JM, Lee N, Perez-Cervantes C, Burnicka-Turek O, Poss KD, Weber CR, de Laat W, Ruthenburg AJ and Moskowitz IP. Transcription-factor-dependent enhancer transcription defines a gene regulatory network for cardiac rhythm. *Elife*. 2017;6.

163. Hogan NT, Whalen MB, Stolze LK, Hadeli NK, Lam MT, Springstead JR, Glass CK and Romanoski CE. Transcriptional networks specifying homeostatic and inflammatory programs of gene expression in human aortic endothelial cells. *Elife*. 2017;6.

164. Cheng Y, Ma Z, Kim BH, Wu W, Cayting P, Boyle AP, Sundaram V, Xing X, Dogan N, Li J, Euskirchen G, Lin S, Lin Y, Visel A, Kawli T, Yang X, Patacsil D, Keller CA, Giardine B, mouse EC, Kundaje A, Wang T, Pennacchio LA, Weng Z, Hardison RC and Snyder MP. Principles of regulatory information conservation between mouse and human. *Nature*. 2014;515:371-375.

165. Blow MJ, McCulley DJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Bristow J, Ren B, Black BL, Rubin EM, Visel A and Pennacchio LA. ChIP-Seq identification of weakly conserved heart enhancers. *Nat Genet*. 2010;42:806-10.

166. May D, Blow MJ, Kaplan T, McCulley DJ, Jensen BC, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Afzal V, Simpson PC, Rubin EM, Black BL, Bristow J,

Pennacchio LA and Visel A. Large-scale discovery of enhancers from human heart tissue. *Nat Genet*. 2012;44:89-93.

167. Das S, Senapati P, Chen Z, Reddy MA, Ganguly R, Lanting L, Mandi V, Bansal A, Leung A, Zhang S, Jia Y, Wu X, Schones DE and Natarajan R. Regulation of angiotensin II actions by enhancers and super-enhancers in vascular smooth muscle cells. *Nature communications*. 2017;8:1467.

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

169. Papait R, Cattaneo P, Kunderfranco P, Greco C, Carullo P, Guffanti A, Vigano V, Stirparo GG, Latronico MV, Hasenfuss G, Chen J and 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.

170. Stratton MS, Lin CY, Anand P, Tatman PD, Ferguson BS, Wickers ST, Ambardekar AV, Sucharov CC, Bradner JE, Haldar SM and McKinsey TA. Signal-Dependent Recruitment of BRD4 to Cardiomyocyte Super-Enhancers Is Suppressed by a MicroRNA. *Cell Rep*. 2016;16:1366-78.

171. Flavahan WA, Gaskell E and Bernstein BE. Epigenetic plasticity and the hallmarks of cancer. *Science*. 2017;357.

172. Li J, Miao L, Shieh D, Spiotto E, Li J, Zhou B, Paul A, Schwartz RJ, Firulli AB, Singer HA, Huang G and Wu M. Single-Cell Lineage Tracing Reveals that Oriented Cell Division Contributes to Trabecular Morphogenesis and Regional Specification. *Cell Rep*. 2016;15:158-70.

173. Meilhac SM, Esner M, Kelly RG, Nicolas JF and Buckingham ME. The clonal origin of myocardial cells in different regions of the embryonic mouse heart. *Developmental cell*. 2004;6:685-98.

174. Moretti A, Caron L, Nakano A, Lam JT, Bernshausen A, Chen Y, Qyang Y, Bu L, Sasaki M, Martin-Puig S, Sun Y, Evans SM, Laugwitz KL and Chien KR. Multipotent embryonic isl1+ progenitor cells lead to cardiac, smooth muscle, and endothelial cell diversification. *Cell*. 2006;127:1151-65.

175. Domian IJ, Chiravuri M, van der Meer P, Feinberg AW, Shi X, Shao Y, Wu SM, Parker KK and Chien KR. Generation of functional ventricular heart muscle from mouse ventricular progenitor cells. *Science*. 2009;326:426-9.

176. Spater D, Hansson EM, Zangi L and Chien KR. How to make a cardiomyocyte. *Development*. 2014;141:4418-31.

177. Kuhn EN and Wu SM. Origin of cardiac progenitor cells in the developing and postnatal heart. *J Cell Physiol*. 2010;225:321-5.

178. DeLaughter DM, Bick AG, Wakimoto H, McKean D, Gorham JM, Kathiriya IS, Hinson JT, Homsy J, Gray J, Pu W, Bruneau BG, Seidman JG and Seidman CE. Single-Cell Resolution of Temporal Gene Expression during Heart Development. *Developmental cell*. 2016;39:480-490.

179. See K, Tan WLW, Lim EH, Tiang Z, Lee LT, Li PYQ, Luu TDA, Ackers-Johnson M and Foo RS. Single cardiomyocyte nuclear transcriptomes reveal a lincRNA-regulated de-differentiation and cell cycle stress-response in vivo. *Nature communications*. 2017;8:225.

180. Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, Chang HY and Greenleaf WJ. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature*. 2015;523:486-90.

181. Wu X, Zhang T, Bossuyt J, Li X, McKinsey TA, Dedman JR, Olson EN, Chen J, Brown JH and Bers DM. Local InsP3-dependent perinuclear Ca2+ signaling in cardiac myocyte excitation-transcription coupling. *J Clin Invest*. 2006;116:675-82.

182. McBrian MA, Behbahan IS, Ferrari R, Su T, Huang TW, Li K, Hong CS, Christofk HR, Vogelauer M, Seligson DB and Kurdistani SK. Histone acetylation regulates intracellular pH. *Mol Cell*. 2013;49:310-21.

183. Simithy J, Sidoli S, Yuan ZF, Coradin M, Bhanu NV, Marchione DM, Klein BJ, Bazilevsky GA, McCullough CE, Magin RS, Kutateladze TG, Snyder NW, Marmorstein R and Garcia BA. Characterization of histone acylations links chromatin modifications with metabolism. *Nature communications*. 2017;8:1141.

184. Worman HJ, Fong LG, Muchir A and Young SG. Laminopathies and the long strange trip from basic cell biology to therapy. *J Clin Invest*. 2009;119:1825-36.

185. Schreiner SM, Koo PK, Zhao Y, Mochrie SG and King MC. The tethering of chromatin to the nuclear envelope supports nuclear mechanics. *Nature communications*. 2015;6:7159.

186. Stephens AD, Liu PZ, Banigan EJ, Almassalha LM, Backman V, Adam SA, Goldman RD and Marko JF. Chromatin histone modifications and rigidity affect nuclear morphology independent of lamins. *Mol Biol Cell*. 2018;29:220-233.

187. Dou Z, Ghosh K, Vizioli MG, Zhu J, Sen P, Wangensteen KJ, Simithy J, Lan Y, Lin Y, Zhou Z, Capell BC, Xu C, Xu M, Kieckhaefer JE, Jiang T, Shoshkes-Carmel M, Tanim

K, Barber GN, Seykora JT, Millar SE, Kaestner KH, Garcia BA, Adams PD and Berger SL. Cytoplasmic chromatin triggers inflammation in senescence and cancer. *Nature*. 2017;550:402-406.

188. Solovei I, Kreysing M, Lanctot C, Kosem S, Peichl L, Cremer T, Guck J and Joffe B. Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. *Cell*. 2009;137:356-68.

189. Edelman LB and Fraser P. Transcription factories: genetic programming in three dimensions. *Curr Opin Genet Dev*. 2012;22:110-4.

190. Sayed D, He M, Yang Z, Lin L and Abdellatif M. Transcriptional Regulation Patterns Revealed by High Resolution Chromatin Immunoprecipitation during Cardiac Hypertrophy. *J Biol Chem*. 2013;288:2546-58.

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

192. Ou HD, Phan S, Deerinck TJ, Thor A, Ellisman MH and O'Shea CC. ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science*. 2017;357.

193. Poleshko A, Shah PP, Gupta M, Babu A, Morley MP, Manderfield LJ, Ifkovits JL, Calderon D, Aghajanian H, Sierra-Pagan JE, Sun Z, Wang Q, Li L, Dubois NC, Morrisey EE, Lazar MA, Smith CL, Epstein JA and Jain R. Genome-Nuclear Lamina Interactions Regulate Cardiac Stem Cell Lineage Restriction. *Cell*. 2017;171:573-587 e14.

194. 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 and 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.

195. Dekker J and Mirny L. The 3D Genome as Moderator of Chromosomal Communication. *Cell*. 2016;164:1110-21.

196. Schmitt AD, Hu M and Ren B. Genome-wide mapping and analysis of chromosome architecture. *Nature reviews Molecular cell biology*. 2016;17:743-755.

197. Davies JO, Oudelaar AM, Higgs DR and Hughes JR. How best to identify chromosomal interactions: a comparison of approaches. *Nat Methods*. 2017;14:125-134.

198. Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA and Bruneau BG. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell*. 2017;169:930-944 e22.

199. Schwarzer W, Abdennur N, Goloborodko A, Pekowska A, Fudenberg G, Loe-Mie Y, Fonseca NA, Huber W, C HH, Mirny L and Spitz F. Two independent modes of chromatin organization revealed by cohesin removal. *Nature*. 2017;551:51-56.

200. Nagano T, Lubling Y, Varnai C, Dudley C, Leung W, Baran Y, Mendelson Cohen N, Wingett S, Fraser P and Tanay A. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature*. 2017;547:61-67.

201. Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, Mirny LA and Dekker J. Organization of the mitotic chromosome. *Science*. 2013;342:948-53.

202. Du Z, Zheng H, Huang B, Ma R, Wu J, Zhang X, He J, Xiang Y, Wang Q, Li Y, Ma J, Zhang X, Zhang K, Wang Y, Zhang MQ, Gao J, Dixon JR, Wang X, Zeng J and Xie W.

Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature*. 2017;547:232-235.

203. Tjong H, Li W, Kalhor R, Dai C, Hao S, Gong K, Zhou Y, Li H, Zhou XJ, Le Gros MA, Larabell CA, Chen L and Alber F. Population-based 3D genome structure analysis reveals driving forces in spatial genome organization. *Proc Natl Acad Sci U S A*. 2016;113:E1663-72.

204. Chen H, Monte E, Parvatiyar MS, Rosa-Garrido M, Franklin S and Vondriska TM. Structural considerations for chromatin state models with transcription as a functional readout. *FEBS letters*. 2012;586:3548-54.

205. Ernst J and Kellis M. Chromatin-state discovery and genome annotation with ChromHMM. *Nature protocols*. 2017;12:2478-2492.

206. Segal E and Widom J. From DNA sequence to transcriptional behaviour: a quantitative approach. *Nat Rev Genet*. 2009;10:443-56.

207. Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X and Karpen GH. Phase separation drives heterochromatin domain formation. *Nature*. 2017;547:241-245.

208. Song K, Nam YJ, Luo X, Qi X, Tan W, Huang GN, Acharya A, Smith CL, Tallquist MD, Neilson EG, Hill JA, Bassel-Duby R and Olson EN. Heart repair by reprogramming non-myocytes with cardiac transcription factors. *Nature*. 2012;485:599-604.

209. Ieda M, Fu JD, Delgado-Olguin P, Vedantham V, Hayashi Y, Bruneau BG and Srivastava D. Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell*. 2010;142:375-86.

210. Liu Z, Chen O, Zheng M, Wang L, Zhou Y, Yin C, Liu J and Qian L. Re-patterning of H3K27me3, H3K4me3 and DNA methylation during fibroblast conversion into induced cardiomyocytes. *Stem cell research*. 2016;16:507-18.

211. Zhao MT, Shao NY, Hu S, Ma N, Srinivasan R, Jahanbani F, Lee J, Zhang SL, Snyder MP and Wu JC. Cell Type-Specific Chromatin Signatures Underline Regulatory DNA Elements in Human Induced Pluripotent Stem Cells and Somatic Cells. *Circ Res*. 2017;121:1237-1250.

212. Tompkins JD, Jung M, Chen CY, Lin Z, Ye J, Godatha S, Lizhar E, Wu X, Hsu D, Couture LA and Riggs AD. Mapping Human Pluripotent-to-Cardiomyocyte Differentiation: Methylomes, Transcriptomes, and Exon DNA Methylation "Memories". *EBioMedicine*. 2016;4:74-85.

213. Zhang Y, Zhong JF, Qiu H, MacLellan WR, Marban E and Wang C. Epigenomic Reprogramming of Adult Cardiomyocyte-Derived Cardiac Progenitor Cells. *Sci Rep*. 2015;5:17686.

214. 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 and Bruneau BG. Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage. *Cell*. 2012;151:206-20.

215. 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 and Murry CE. A temporal chromatin signature in human embryonic stem cells identifies regulators of cardiac development. *Cell*. 2012;151:221-32.

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

217. Monte E, Rosa-Garrido M, Karbassi E, Chen H, Lopez R, Rau CD, Wang J, Nelson SF, Wu Y, Stefani E, Lusis AJ, Wang Y, Kurdistani SK, Franklin S and Vondriska TM. Reciprocal Regulation of the Cardiac Epigenome by Chromatin Structural Proteins Hmgb and Ctcf: IMPLICATIONS FOR TRANSCRIPTIONAL REGULATION. *J Biol Chem*. 2016;291:15428-46.

218. Smith JG, Felix JF, Morrison AC, Kalogeropoulos A, Trompet S, Wilk JB, Gidlof O, Wang X, Morley M, Mendelson M, Joehanes R, Ligthart S, Shan X, Bis JC, Wang YA, Sjogren M, Ngwa J, Brandimarto J, Stott DJ, Aguilar D, Rice KM, Sesso HD, Demissie S, Buckley BM, Taylor KD, Ford I, Yao C, Liu C, consortium C-S, EchoGen c, consortium Q-I, consortium C-Q, Sotoodehnia N, van der Harst P, Stricker BH, Kritchevsky SB, Liu Y, Gaziano JM, Hofman A, Moravec CS, Uitterlinden AG, Kellis M, van Meurs JB, Margulies KB, Dehghan A, Levy D, Olde B, Psaty BM, Cupples LA, Jukema JW, Djousse L, Franco OH, Boerwinkle E, Boyer LA, Newton-Cheh C, Butler J, Vasan RS, Cappola TP and Smith NL. Discovery of Genetic Variation on Chromosome 5q22 Associated with Mortality in Heart Failure. *PLoS genetics*. 2016;12:e1006034.

219. 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 and Posch AE. Epigenome-wide association

study identifies cardiac gene patterning and a novel class of biomarkers for heart failure. *Circulation*. 2017;136:1528-1544.

220. Bohacek J and Mansuy IM. Molecular insights into transgenerational non-genetic inheritance of acquired behaviours. *Nat Rev Genet*. 2015;16:641-52.

221. Turner BM. Environmental sensing by chromatin: An epigenetic contribution to evolutionary change. *FEBS letters*. 2011.

222. Ferguson JF, Allayee H, Gerszten RE, Ideraabdullah F, Kris-Etherton PM, Ordovas JM, Rimm EB, Wang TJ, Bennett BJ, American Heart Association Council on Functional G, Translational Biology CoE, Prevention and Stroke C. Nutrigenomics, the Microbiome, and Gene-Environment Interactions: New Directions in Cardiovascular Disease Research, Prevention, and Treatment: A Scientific Statement From the American Heart Association. *Circulation Cardiovascular genetics*. 2016;9:291-313.

223. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, Reese SE, Markunas CA, Richmond RC, Xu CJ, Kupers LK, Oh SS, Hoyo C, Gruzieva O, Soderhall C, Salas LA, Baiz N, Zhang H, Lepeule J, Ruiz C, Ligthart S, Wang T, Taylor JA, Duijts L, Sharp GC, Jankipersadsing SA, Nilsen RM, Vaez A, Fallin MD, Hu D, Litonjua AA, Fuemmeler BF, Huen K, Kere J, Kull I, Munthe-Kaas MC, Gehring U, Bustamante M, Saurel-Coubizolles MJ, Quraishi BM, Ren J, Tost J, Gonzalez JR, Peters MJ, Haberg SE, Xu Z, van Meurs JB, Gaunt TR, Kerkhof M, Corpeleijn E, Feinberg AP, Eng C, Baccarelli AA, Benjamin Neelon SE, Bradman A, Merid SK, Bergstrom A, Herceg Z, Hernandez-Vargas H, Brunekreef B, Pinart M, Heude B, Ewart S, Yao J, Lemonnier N, Franco OH, Wu MC, Hofman A, McArdle W, Van der Vlies P, Falahi F, Gillman MW, Barcellos LF, Kumar A, Wickman M, Guerra S, Charles MA, Holloway J, Auffray C, Tiemeier HW, Smith

GD, Postma D, Hivert MF, Eskenazi B, Vrijheid M, Arshad H, Anto JM, Dehghan A, Karmaus W, Annesi-Maesano I, Sunyer J, Ghantous A, Pershagen G, Holland N, Murphy SK, DeMeo DL, Burchard EG, Ladd-Acosta C, Snieder H, Nystad W, Koppelman GH, Relton CL, Jaddoe VW, Wilcox A, Melen E and London SJ. DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *American journal of human genetics*. 2016;98:680-96.

224. Patterson AJ, Xiao D, Xiong F, Dixon B and Zhang L. Hypoxia-derived oxidative stress mediates epigenetic repression of PKCepsilon gene in foetal rat hearts. *Cardiovasc Res*. 2012;93:302-10.

225. Huypens P, Sass S, Wu M, Dyckhoff D, Tschop M, Theis F, Marschall S, Hrabe de Angelis M and Beckers J. Epigenetic germline inheritance of diet-induced obesity and insulin resistance. *Nat Genet*. 2016;48:497-9.

226. Carone BR, Fauquier L, Habib N, Shea JM, Hart CE, Li R, Bock C, Li C, Gu H, Zamore PD, Meissner A, Weng Z, Hofmann HA, Friedman N and Rando OJ. Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. *Cell*. 2010;143:1084-96.

227. Leung A, Parks BW, Du J, Trac C, Setten R, Chen Y, Brown K, Lusis AJ, Natarajan R and Schones DE. Open chromatin profiling in mice livers reveals unique chromatin variations induced by high fat diet. *J Biol Chem*. 2014;289:23557-67.

228. Xu H, Wang F, Kranzler HR, Gelernter J and Zhang H. Alcohol and nicotine codependence-associated DNA methylation changes in promoter regions of addiction-related genes. *Sci Rep*. 2017;7:41816.

229. Wang JJ, Aboulhosn JA, Hofer IS, Mahajan A, Wang Y and Vondriska TM. Operationalizing Precision Cardiovascular Medicine: Three Innovations. *Circ Res*. 2016;119:984-987.

230. MacRae CA, Roden DM and Loscalzo J. The Future of Cardiovascular Therapeutics. *Circulation*. 2016;133:2610-7.

231. Liu XS, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, Shu J, Dadon D, Young RA and Jaenisch R. Editing DNA Methylation in the Mammalian Genome. *Cell*. 2016;167:233-247 e17.

232. Morgan SL, Mariano NC, Bermudez A, Arruda NL, Wu F, Luo Y, Shankar G, Jia L, Chen H, Hu JF, Hoffman AR, Huang CC, Pitteri SJ and Wang KC. Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. *Nature communications*. 2017;8:15993.

Chapter 2: High resolution mapping of chromatin conformation in cardiac myocytes reveals structural remodeling of the epigenome in heart failure

Manuel Rosa-Garrido, PhD, Douglas J. Chapski, BS, Anthony D. Schmitt, PhD, Todd H. Kimball, BS, Elaheh Karbassi, PhD, Emma Monte, PhD, Enrique Balderas, PhD, Matteo Pellegrini, PhD, Tsai-Ting Shih, BS, Elizabeth Soehalim, BS, David Liem, MD, PhD, Peipei Ping, PhD, Niels J. Galjart, PhD, Shuxun Ren, MD, Yibin Wang, PhD, Bing Ren, PhD, Thomas M. Vondriska PhD

[This research was originally published by Rosa-Garrido *et al.* High resolution mapping of chromatin conformation in cardiac myocytes reveals structural remodeling of the epigenome in heart failure. *Circulation*. 2017 Oct 24;136(17):1613-1625. PMID: 28802249 © The Authors.]

Abstract

Background: Cardiovascular disease is associated with epigenomic changes in the heart, however the endogenous structure of cardiac myocyte chromatin has never been determined.

Methods: To investigate the mechanisms of epigenomic function in the heart, genome-wide chromatin conformation capture (Hi-C) and DNA sequencing were performed in adult cardiac myocytes following development of pressure overload-induced hypertrophy. Mice with cardiac-specific deletion of CTCF (a ubiquitous chromatin structural protein) were generated to explore the role of this protein in chromatin structure and cardiac

phenotype. Transcriptome analyses by RNA-seq were conducted as a functional readout of the epigenomic structural changes.

Results: Depletion of CTCF was sufficient to induce heart failure in mice and human heart failure patients receiving mechanical unloading via left ventricular assist devices show increased CTCF abundance. Chromatin structural analyses revealed interactions within the cardiac myocyte genome at 5kb resolution, enabling examination of intra- and inter-chromosomal events, and providing a resource for future cardiac epigenomic investigations. Pressure overload or CTCF depletion selectively altered boundary strength between topologically associating domains and A/B compartmentalization, measurements of genome accessibility. Heart failure involved decreased stability of chromatin interactions around disease-causing genes. In addition, pressure overload or CTCF depletion remodeled long-range interactions of cardiac enhancers, resulting in a significant decrease in local chromatin interactions around these functional elements.

Conclusions: These findings provide a high-resolution chromatin architecture resource for cardiac epigenomic investigations and demonstrate that global structural remodeling of chromatin underpins heart failure. The newly identified principles of endogenous chromatin structure have key implications for epigenetic therapy.

KEY WORDS

Epigenetics, genomics, heart failure, hypertrophy, CTCF, chromatin conformation capture

CLINICAL PERSPECTIVE

What Is New?

- Chromatin capture and DNA sequencing were used to determine the endogenous structure of the cardiac myocyte epigenome.
- Physical interactions between regulatory elements and cardiac disease genes have been determined in basal and disease settings.
- The role of the chromatin structural protein CTCF was examined using an *in vivo* loss-of-function model, revealing its role in chromatin organization and disease.

What Are the Clinical Implications?

- Epigenomic plasticity is identified as a common feature of cardiac pathophysiology induced by distinct stimuli.
- Knowledge of the dynamics of genomic interactions in disease may enable new strategies for therapeutic intervention.

INTRODUCTION

Heart failure is a devastating condition that affects 6.5 million adults in the USA.[1] Although it is a multi-organ disease, heart failure is driven by changes in cardiac myocyte biology, including cell death, calcium handling, myofilament function, metabolism and other factors. Underlying this complex cellular malfunction are gene expression changes, orchestrated by a network of transcription factors and chromatin remodeling enzymes.[2-5] Despite this knowledge, the basic folding principles of the cardiac myocyte epigenome have never been revealed and the role of chromatin structural changes in cardiovascular disease is unknown.

Correct packaging of the genome within the nucleus determines appropriate gene expression and cellular function.[6] Nucleosomes are differentially positioned along chromosomes, a process controlled by chromatin remodelers and histone modifying enzymes.[7] Tracks of nucleosomes adopt fiber-like structures that in turn compose the three-dimensional architecture of endogenous chromatin, a process that is necessary for orderly storage and controlled accessibility of genetic information.[8]

Chromatin modifications are engaged in a developmentally tuned manner to enable both the unidirectional procession of differentiation and transcriptome stability. Cells employ enhancers[9]—distal regulatory regions that host histone post-translational modifications and have an increasingly appreciated role in cardiovascular disease[10-12]— and chromatin binding proteins which, together with transcription factors, sculpt the transcriptome. Among chromatin structural proteins, CTCF has been attributed a key role in modulating genome architecture[13] and maintaining regions of genome accessibility across cell types.[14] Once cellular lineages have been established, prevailing proteomic

programs ensure that upon division, the correct epigenomic landscape is reestablished in daughter cells, maintaining lineage fidelity. An underexplored area of chromatin biology is how epigenomic stability is maintained *in vivo* when cells have exited the cell cycle. Adult mammalian cardiac myocytes do not readily divide and the heart lacks robust regenerative capacity. While epigenetic transitions in cardiovascular development have been investigated in cell culture,[15, 16] much less is known about the adult heart. Previous studies have demonstrated a role for histone modifying enzymes[3] and the readers of these modifications,[17] but whether three-dimensional configuration of the genome contributes to heart disease remains to be determined.

Recent investigations have demonstrated the existence of topologically associating domains (TADs) as regions of chromatin that exhibit a higher level of local interactions. These structural neighborhoods of the genome may facilitate co-regulation of gene expression and are shared across many biological conditions, suggesting they are fundamental features of the genome.[8, 18] However the specific nature of interactions within and between these TADs in cardiac cells, and the additional hierarchical levels of chromatin packaging (including local interactions, long range interactions, enhancer-gene interactions and chromatin looping) utilized in the cardiac myocyte nucleus in health and disease, are unknown.

The findings of the present study establish a high-resolution, genome-wide resource of endogenous chromatin interactions in cardiac myocytes, which can be used by future investigations to examine virtually any genomic locus. In addition, the results demonstrate specific areas of the epigenome that are structurally reorganized in pressure overload-induced heart disease, revealing how enhancers interact with disease causing

genes and the role of chromatin looping in cardiac gene expression. Using a physiologically relevant form of afterload-induced cardiac disease (plus a lineage targeted loss-of-function CTCF knockout mouse as a unique model for comparison), this investigation examines the role of global epigenome structure in heart failure.

METHODS

All animal studies were approved by the UCLA Animal Research Committee in compliance with the NIH Guide for the Care and Use of Laboratory Animals. Human samples used in this study were procured in the Ronald Reagan Medical Center at UCLA following patient consent according to Institutional Review Board-approved protocol. This study used DNA from adult mouse cardiac myocytes isolated from control mice or two distinct disease models (pressure overload-induced hypertrophy and CTCF-KO) to perform chromosome conformation capture (Hi-C) analysis to determine endogenous chromatin structure (all sequencing-based studies were performed on $n \geq 3$ biological replicates per group; n values for all other endpoints are provided in figure legends). Transcriptome analyses were carried out using RNA-seq to examine the functional readouts of any structural changes in the epigenome. Chromatin structure was then examined at multiple scales to determine structural units, hierarchical organization and three-dimensional regulation of accessibility and transcription.

Statistical Analyses

All sequencing-based studies were performed on $n \geq 3$ biological replicates per group; n values for all other endpoints are provided in figure legends. The statistical tests performed for each endpoint are listed in the individual figure legends, described in the Online Supplemental Methods and are summarized here: Student's t-test: Figure 2-1a, 2-1e, 2-8c, 2-8e and 2-10; Tukey test: Figure 2-1c; Wilcoxon test: Figure 2-2f; Pearson correlation: Figure 2-9b; RNA-seq data: Cuffdiff was used to determine differential gene expression and assign p-values and q-values for each gene; for interaction analyses in

Figures 2-4b-d and Figure 2-5: Fit-Hi-C was used to filter for *cis* interactions with *q*-value < 0.01 which were then used to map interactions.

Detailed methodology for all experiments and analyses are available in the Online Supplement.

RESULTS

Chromosome conformation capture

Chromatin was purified from isolated cardiac myocytes and subjected to chromosome conformation capture (Hi-C). Mappability of reads was 63-65% across experimental groups (Figure 2-7A) at a depth of sequencing enabling examination of interactions with 5kb resolution (Figure 2-7B). No significant differences were observed across experimental groups in terms of number of interactions analyzed and run-to-run variability was low (Figure 2-7C). Unless otherwise noted in the figures, all "control" groups were from myocytes isolated from untreated adult wild type mice (all mice used in this study are C57BL/6 background). With this high-resolution resource, one can explore the physical environment of virtually any genomic locus in the cardiac myocyte at multiple scales including topologically associating domains, A/B compartmentalization, chromatin looping, enhancer interactions and gene interactions. Sequencing data have been uploaded to NCBI with the GEO accession number GSE96693 and are publically available.

Loss of CTCF causes cardiomyopathy

A mouse line with inducible, cardiac myocyte-specific ablation of CTCF was generated. Mice with a *loxp* flanked *Ctcf* allele ($Ctcf^{flox/flox}$ mice) were crossed with mice expressing a transgenic tamoxifen-inducible Cre protein under the control of the α-MHC promoter (MerCreMer mice). When administered tamoxifen in the diet for 5 weeks followed by 1 week on regular chow (Figure 2-8A, right panel), $Ctcf^{flox/flox}$-$MerCreMer^{+/-}$ mice display selective excision of the targeted region of the *Ctcf* gene (Figure 2-8B) and

exhibit a gradual loss of the transcript in isolated cardiac myocytes (Figure 2-8C). Further depletion in CTCF levels was associated with poor survival (Figure 2-8D); as a control, treatment of $Ctcf^{wt/wt}$-$MerCreMer^{+/-}$ mice with tamoxifen had no sustained effect on cardiac function or morphology (Figure 2-8G-I). The experimental mice—$Ctcf^{flox/flox}$-$MerCreMer^{+/-}$ animals treated with tamoxifen to induce CTCF depletion—had ~80% reduction in CTCF protein levels (Figure 2-8F) and are henceforth referred to as CTCF-KO mice. A separate cohort of wild type mice were subjected to pressure overload stress (through surgical application of transverse aortic constriction,[19] abbreviated TAC) which induces heart failure through clinically-relevant pathophysiological adaptation.

Phenotypic examination of the CTCF-KO mice revealed a striking cardiomyopathy (Figure 2-1 and Figure 2-8G-I). Loss of CTCF leads to impaired ejection fraction, left ventricular chamber dilation and muscle hypertrophy at the organ and cell level (Figure 2-1A-C). The TAC mice, by contrast, also exhibit hypertrophy accompanied by more modest changes in both chamber size and ejection fraction, demonstrating that these animals (CTCF-KO and TAC mice) represent distinct pathophysiological models. Examining CTCF expression in a genetically heterogeneous population of wild type mice revealed consistent down-regulation following pathologic stimuli and an inverse correlation with pathologic measurements of heart size (Supplemental Figure 2-9A and B; interestingly, CTCF protein and mRNA levels were unchanged after TAC [*data not shown*], further suggesting that these models of disease, although they share some phenotypic characteristics, have important molecular and pathophysiological distinctions). The time course for chamber dilation, diminished EF and CTCF depletion in CTCF-KO mice were similar (Figure 2-8C and 2-8H), and agreed with that for pathologic

gene activation (Figure 2-10). CTCF-KO and TAC led to fibrotic deposition (Figure 2-1B, bottom panel) and activation of known heart failure genes[2] (Figure 2-1D; Figure 2-7D shows RNA-seq data quality statistics). Caspase-3 cleavage assays to measure apoptosis showed no difference between CTCF-KO and control animals (Figure 2-11, suggesting, but not proving, that impairment of cardiac function is secondary to necrosis). In human heart tissues harvested before and after implantation of left ventricular assist devices (LVADs)—which mechanically unload the heart, allowing reverse remodeling of diseased tissue—CTCF levels were increased after device implantation (Figure 2-1E-G; see Figure 2-12 for patient clinical data; unavailability of healthy human hearts precluded measurement of CTCF expression in these tissues).

Endogenous chromatin architecture in healthy and diseased myocytes

We next sought to examine the large-scale alterations in chromatin structure underlying heart disease in the mice subjected to TAC. There are several important differences in the pathophysiology between the CTCF-KO and TAC mice, and thus the former is used herein as a type of alternate disease model to investigate similarities and differences in global chromatin changes in different forms of heart failure. Analyses of topologically associating domain (TAD) architecture in wild type, CTCF-KO mice or TAC mice demonstrate TADs to be stable features of chromatin structure (Figure 2-2A; average TAD size = 445kb). The number of TAD boundaries across the entire genome varied by <2%, with 3686, 3746 and 3705 boundaries called in control, CTCF-KO and pressure overload, respectively. Boundary strength, however, was differentially altered across the genome (Figure 2-2B shows the distribution across chromosomes for control

versus TAC; Figure 2-2C shows quantitation; Figure 2-13A shows all chromosomes in all conditions; Figure 2-14A shows PCA for boundary strength). CTCF-KO or TAC chromatin each hosted new boundaries (Figure 2-14B; "new" include shifted boundaries and *de novo* formation; note: local rate of change in insulation score determines boundary strength[20]).

Chromatin can be divided into active and inactive regions,[21] called A and B compartments respectively, which were affected to only a minor degree by loss of CTCF or TAC. Compartmentalization changes were sparse (Figure 2-2D-E) and ranged from <2 to ~8% of interaction bins genome wide. The scale of the total genome changing compartmentalization was almost identical (~4%) after CTCF-KO or TAC (Figure 2-2E). Changes in compartmentalization correlated positively with gene expression: that is, genes moving from A to B were down-regulated and B to A up-regulated (Figure 2-2F). Loss of CTCF elicited a transcriptional profile that shared some (41% of genes) features with TAC (Figures 2-2G-H; Figure 2-14C shows PCA of transcriptome data) as well as known marker gene activation (Figure 2-1D). Genes determined by ChIP-seq to harbor CTCF in their transcription start sites (TSSs) were enriched in pathways associated with cardiac disease (Figure 2-14D).

Chromatin looping is altered with CTCF KO or TAC

The hierarchical nature of chromatin results in a preponderance of short-range interactions, while simultaneously necessitating long-range looping (Figure 2-3A). One such long-range interaction affected by CTCF depletion and TAC is shown in Figure 2-3B. Quantitative analysis shows that loss of CTCF or TAC were associated with a

decrease in the total number of long-range loops, without a gross change in loop size (Figure 2-3C). Of dynamic loops in disease, 51% of those disappearing were shared between CTCF-KO and TAC, whereas only 15% of those appearing were shared; the raw number of loops lost in disease was also greater than the number formed (Figure 2-3D). In control chromatin, 37% (3056 out of 8240; Figure 2-3E) of loops were flanked by two CTCF peaks (78% ≥1 peak) and loss of one CTCF peak was sufficient to destroy loops in 37% of the cases (326 out of 879; Figure 2-3F). Within reorganized loops, enrichment for genes in pathways associated with cardiovascular function was observed (Figure 2-14E; a caveat here is that analysis of cardiac myocytes may increase the chances that enriched ontology terms are of cardiovascular relevance).

Among total loops, a subset is responsible for bringing together enhancers and promoters (Figure 2-3G-H). In agreement with aberrant gene expression in both conditions (Figure 2-2F-H), lost loops in CTCF-KO and TAC chromatin were enriched in genes associated with cardiac pathology (Figure 2-14F).

Heart failure involves altered enhancer-gene interactions

Focusing more deeply on significant interactions that contribute to gene expression can reveal the physiological implications of genomic structure. For example, *Ppp3ca*, a gene implicated in calcium dependent signaling and cardiac hypertrophy[4] exhibited a consistent decrease in total interactions in CTCF-KO or TAC conditions (Figure 2-4A).

Notably, of 9194 expressed genes for which significant (q<0.01) Fit-Hi-C data were available, 60% had the same direction of change of interactions in CTCF-KO and TAC

conditions. Of the 40% of genes with unique behavior in the diseased condition, the majority (75%) were different because one of the two (CTCF-KO or TAC) exhibited no change in interaction; 25% were situations in which the change in interaction was opposite between CTCF-KO and TAC. Of the genes with shared changes in interactions in CTCF-KO and TAC that also underwent differential expression (3651 out of 5443), the vast majority (86%) experienced decreased local interactions (Figure 2-4B). These findings support a trend in which more fluid chromatin interactions are associated with gene expression and where alterations in expression (when they occur) after perturbation are more likely to be in the same direction in CTCF-KO and TAC compared to control. Figure 2-4C shows the top 40 differentially expressed genes selected for the number of interactions in the control condition, demonstrating a consistent decrease in interactions in CTCF-KO and TAC.

Distinct mechanisms of gene regulation were associated with distinct changes in chromatin microenvironment, represented in three example genes (Figure 2-4D). Expression of *Nppa* (Figure 2-4D, left; top is control interaction matrix, middle has specific interactions colored by significance and bottom has RNA-seq expression tracks) a known marker for heart failure in mice and humans,[2] was increased in CTCF-KO and TAC. Nearby local interactions are dense but relatively unchanging, whereas those impinging on the *Nppa* gene (demarcated by vertical lines) are decreased. Expression of *Kcnd2*, a potassium channel implicated in sudden death,[22] was down-regulated and the locus exhibited a notable decrease in long-range interactions in CTCF-KO and TAC (Figure 2-4D, center). Lastly, a third class of regulatory scheme is represented by *Mef2c*, whose expression changes with neither CTCF-KO nor TAC but which undergoes a high degree

of splicing regulation[23] and is a central driver of cardiac gene expression[24] (Figure 2-4D, right). This locus underwent a consistent decrease in chromatin interactions up- and down-stream of the gene in CTCF-KO and TAC, suggesting a loosening of the local chromatin environment.

Enhancers are regions of the genome that positively influence gene expression and are identified on the basis of diagnostic histone post-translational modifications. How enhancers are regulated in the three-dimensional context of the cardiac nucleus is unknown. Of previously identified enhancers from adult cardiac myocytes,[11] 5050 contained significant Fit-Hi-C interactions. These enhancers consistently exhibited decreased interactions after perturbation (Figure 2-5A). Indeed 47% of enhancers exhibited decreased interactions in CTCF-KO and 67% of enhancers exhibited decreased interactions in TAC (Figure 2-5B, left and center). Remarkably, of the enhancers that were shared between the two conditions, 85% showed a decrease of interactions in both CTCF-KO and TAC (Figure 2-5B, right). Figure 2-5C further illustrates that CTCF-KO and TAC were both associated with decreased enhancer-gene interactions (the directionality of gene expression change is reported in the different cohorts of enhancers from Figure 2-5B) and the affected genes were enriched in cardiac pathology pathways (Figure 2-14G). *Rock2*, previously implicated in cardiac disease,[25] is an example gene whose interactions with multiple enhancers are depleted in the setting of CTCF-KO or TAC (Figure 2-5D; Online Tables 1-3). These findings indicate that a global increase in chromatin fluidity around enhancers is a shared feature of cardiac pathology.

DISCUSSION

This study provides a resource of endogenous chromatin architecture in cardiac myocytes and describes the global changes in chromatin interactions during heart failure. With this data, the basal microenvironment of any genomic locus can be explored in the future and changes in this microenvironment can be examined in the setting of pressure overload-induced heart failure. As a model of heart disease caused by increased afterload (including hypertension and aortic stenosis), TAC was used in this study to investigate how global changes in the epigenome participate in disease pathogenesis—the CTCF-KO mouse was used as a complementary tool to investigate the role of a critical chromatin structural protein. Interestingly, loss of CTCF or pressure overload-induced heart failure caused greater dynamics of endogenous chromatin structure and specific reorganization of enhancer-gene interactions. Depletion of CTCF resulted in a phenotype that shared some features of that induced by pressure overload (including fibrosis, hypertrophy, changes in EF and changes in chamber dimensions), although there were notable differences between these phenotypes (particularly in the extent of LV dilation), meaning that CTCF-KO does not fully recapitulate and is thus not a model of pressure overload hypertrophy, but rather, is a form of dilated cardiomyopathy with hypertrophy and fibrosis. Human heart failure patients treated with left ventricular assist devices to unload the heart demonstrated increased CTCF protein and mRNA levels. While this observation is interesting, caution is warranted in the extrapolation of these findings to the clinical arena: future cohort studies in cardiomyopathies of distinct etiology will be required to fully understand the role of CTCF in human heart failure.

Although the pattern of chromosome organization during mitosis is well known, how the genome is structured during interphase or in cells that have exited the cell cycle has remained a mystery until very recently. Interphase nuclei form chromosomal territories (features mainly characterized by microscopy),[26] which serve to segregate entire chromosomes apart from each other. The development of genome-wide DNA sequencing-based approaches in the past few years has enabled precise determination of sub-chromosomal genome architecture.[8, 18, 27] This information—that is, how the genome is packaged in different cells—is important for basic science reasons (e.g. to answer fundamental questions about the structure-function relationship between an invariant genome and a highly variable transcriptome [amongst different cells in a multi-cellular organism]) as well as for translational implications (e.g. many different diseases involve changes in epigenetic machinery, but understanding how these chromatin modifiers exert their control of disease phenotype requires determination of the substrate on which they operate, i.e. the structural conformation of the epigenome). With these considerations in mind, the present study investigates each of the various levels of hierarchical organization of cardiac chromatin including: the formation of TADs and the compartmentalization of large swaths of the genome (Figure 2-2 and Figure 2-13), intermediate structural features, including chromatin loops of various lengths (Figure 2-3), precise remodeling of interactions around genes involved in disease pathogenesis (Figure 2-4), and the structural organization of non-coding functional units (i.e. enhancers, which incorporates existing knowledge of post-translational modifications decorating histones at specific genomic loci) near the genes they modify (Figure 2-5).

The present study and previous work in the field suggest that heart failure is associated with a plastic epigenome (Figure 2-6): widespread changes in boundary strength (Figure 2-2C); decreased formation of chromatin loops (Figure 2-3C-D); decreased interactions in local chromatin environment of differentially expressed genes (Figure 2-4); and more fluid interactions between enhancers and genes (Figure 2-5). These findings support previous studies of heart failure showing global DNA demethylation,[28] altered histone stoichiometry and a decrease in heterochromatic post-translational modifications,[19, 29, 30] activation of pathogenic noncoding RNAs,[31, 32] and aberrant transcriptional activation.[3, 4, 33] Heart disease has been speculated to be the result of a reversion to more primitive gene expression programs[2]—the present findings extend this hypothesis, identifying shared epigenomic features in response to cellular stress.

While CTCF has been associated with TAD boundaries,[34-36] its structural role has been unclear in post-mitotic cells *in vivo*. The present study explores the role of CTCF in cardiac epigenome stability, but cannot unequivocally conclude whether CTCF is dispensable for TAD maintenance, insofar as residual CTCF was present. A recent investigation, using an auxin-inducible degron approach to deplete CTCF protein in mouse embryonic stem cells, demonstrated that near complete depletion of CTCF was required to influence gross chromatin organization, with TAD structure still observed with as low as 4% of normal CTCF protein levels.[37] In the present paper, because many loops did not harbor CTCF, and a large number of loops with unchanged CTCF binding were altered in their conformation, these findings support a model in which other factors stabilize chromatin looping and CTCF participates in gene regulation through

mechanisms in addition to anchoring. Other insulator proteins (*e.g.* cohesins, which behave similarly to CTCF across a population of mice and are modestly down-regulated in CTCF-KO mice [Figure 2-9A], suggesting that cohesins and CTCF are co-regulated in circumstances of global chromatin remodeling) may stabilize TADs, in addition to histone H3 lysine 27 trimethylation (globally unchanged in CTCF-KO, Figure 2-8F) and perhaps DNA methylation. CTCF is essential for development beyond embryo implantation[38] and longer-term depletion in the heart was lethal (Figure 2-8D). This investigation demonstrates that subtle perturbations in the structure of gene expression domains can result in catastrophic malfunction of the myocyte and heart, in agreement with previous findings implicating CTCF in disease-related transcription in other organs.[39] Cell death in the heart is an active area of investigation[40] and while our studies found no evidence for apoptotic cell death, it is possible, given the progressive nature of this phenotype, that the CTCF-KO cardiomyopathy is driven by one or more forms of cell death to be determined in future studies. Our previous investigations[41] showed that CTCF knockdown was sufficient to induce cell death in HeLa cells but not in neonatal rat ventricular myocytes and other investigations in non-cardiac cells have suggested that aberrant CTCF function leads to abnormal cell death decisions, such as in cancer.[42]

Recent studies have identified a role for loops in the formation of chromatin neighborhoods to insulate against aberrant—and to coordinate synchronized—expression of genes.[43] In addition, ChIP-seq investigations of histone marks have provided important insights into how these features correlate with gene expression.[11] With high resolution chromatin conformation data, the more biologically relevant approach can be taken to examine enhancers in the circumambient context in which they operate.[44]

The present findings are a resource for examining gene regulation by virtually any known, or to-be-discovered, cardiac enhancer as well as to explore newly identified three-dimensional features of the cardiac nucleus.

In summary, these studies point to a common entropic destination for pathologically disturbed chromatin, supporting a model wherein chromatin structural abnormalities underpin the complex cellular networks that go awry in disease and providing a new conceptual framework to engineer therapies.

Acknowledgements:

We appreciate help from Ferhat Ay and Ruyu Tan (La Jolla Institute for Allergy and Immunology) with Fit-Hi-C, as well as Yunjiang Qiu (Bing Ren lab) with Juicer. We thank Aldons J. Lusis and Christoph D. Rau (UCLA) for discussions and assistance. We thank the UCLA Center for Genomics for RNA-seq assistance.

Funding Sources:

MRG was supported by an American Heart Association Post-doctoral Fellowship (16POST27780019); DJC was supported by the Eugene V. Cota-Robles Predoctoral Fellowship and a Ruth L. Kirschstein Institutional Training Grant (T32HL69766); EK and EM were supported by American Heart Association Pre-doctoral Fellowships. This project was supported by the UCLA Cardiovascular Theme and NIH grants HL 105699 (TMV), HL 115238 (TMV), and HL 129639 (YW and TMV).

Author contributions:

Conceptualization, MRG, TMV; Methodology, MRG, DJC, ADS, EB, SR, MP, BR, TMV; Experimentation, MRG, DJC, EB, TS, ES, EK, THK, SR; Data Analysis, MRG, DJC, ADS, EM, BR, TMV; Writing, MRG, DJC, TMV; Revision and Editing, all authors; Resources, NJG, DL, PP, BR, YW, TMV; Funding Acquisition, MRG, YW, TMV.

Disclosures:

None

Supplemental Material

High resolution mapping of chromatin conformation in cardiac myocytes reveals structural remodeling of the epigenome in heart failure

Rosa-Garrido *et al*.

This section contains the extended Materials & Methods Section. Online Tables 1-3, containing enhancer-gene data, are uploaded on the journal website. Sequencing data used in this paper have been deposited in NCBI with GEO accession number GSE96693.

MATERIALS & METHODS

Hi-C Methodology

For each replicate (n≥3 for all sequencing experiments in this paper), one million isolated adult cardiomyocytes were fixed in 1% formaldehyde and underwent *in situ* Hi-C as described,[45] using *MboI* as the restriction enzyme. Hi-C libraries were sequenced on an Illumina HiSeq 4000 instrument. Please see Figure 2-7 for detailed data on read counts, cis/trans interactions, and mappability. All sequencing data have been uploaded to NCBI with the GEO accession number GSE96693.

Bioinformatics Analyses of Hi-C Data

Raw paired-end 50bp reads were mapped to mm10 using BWA-MEM[46] as recently described.[14] Raw genome-wide contact matrices were generated using a custom Perl script. To normalize genome-wide raw contact matrices, we first removed bin pairs that fell outside the 5th and 99.5th percentiles (that had abnormally low or high coverage, respectively). We then used a custom Perl script to perform iterative correction and eigenvalue decomposition (ICE) on the filtered matrices for 50 iterations.[47] We then divided all entries in each genome-wide contact matrix by the corresponding matrix sum, and then multiplied them by the arbitrary value of 1 billion for 5kb resolution or 1 million for 40kb and 50kb resolution.[20] For the main analyses in this study, all biological replicates were combined.

For the individual replicate analyses, we generated replicate-specific contact matrices using the methods described above, and then calculated a Pearson correlation

for the ice-corrected and normalized *cis* bin pairs for which we measured interactions in each replicate pair. We then plotted $\log_{10}$(normalized interactions) for each replicate on a respective axis using the smoothScatter() function in R. Once we determined a high correlation between replicates, we pooled the 3 replicates of each sample type together.

To generate 40kb z-score matrices, we first generated 40kb normalized contact matrices from the combined replicates for each sample type, using the methodologies described above. Z-score *cis* matrices were then generated using the loess method,[20] using the following function from the HiTC package[48] in Bioconductor: normPerExpected(matrix, method="loess", span=0.005, stdev=T). Then, *q*-values were calculated using the zscore2pval.R script in https://github.com/dekkerlab/cworld-dekker. Significant *q*-values < 0.01 were used for downstream analyses and Circos plot generation in R.

Identification of AB compartments at 5kb resolution was performed using the prcomp() function in R and correlating PC1 values to gene density. To generate a heatmap of difference in AB compartmentalization, we first determined the bins on autosomes that underwent compartment change in at least one sample. Then, we performed k-means clustering with $k = 5$. The AB difference graphs were generated using the following logic: If the PC1 value of a given bin of a diseased sample became more positive or less negative when compared to control, then the difference was deemed "positive" and the magnitude was shown in orange as a positive difference. Contrastingly, if the PC1 value of a given bin of a diseased sample became more negative or less positive, then the difference was deemed "negative" and the magnitude was shown in purple as a negative difference. To compare AB compartment changes with gene

expression differences, we determined the AB compartmentalization for the TSS bin for each autosomal gene and then generated boxplots showing compartment change on the x-axis and $\log_2$(fold-change) of FPKM on the y-axis.

Insulation scores were calculated using the method,[20] using a 500kb sliding window across each 5kb normalized *cis* contact matrix. Insulation scores for all samples were normalized together using the normalize.quantiles() function of the preprocessCore package in Bioconductor (originally described[49]). The delta vector for each insulation score profile was generated using a custom R script, using a 500kb sliding window across each 5kb normalized contact matrix, and a 100kb outward distance to the left and right of each bin to calculate the slope of the insulation score. Insulation boundaries were defined as local minima in the insulation score, and boundary strength was calculated. We discarded boundaries with strength < 0.1.[20] Insulation boundary heatmaps were generated by first calculating the insulation boundaries that overlap between control and CTCF knockout, and between control and TAC. Then, we generated heatmaps of overlapping boundary strength changes per autosome.

To call and visualize chromatin loops, we used Juicer[50] and Juicebox.[51] We first ran juicebox_tools_7.0 pre to generate .hic files for each condition, using reads with mapping quality ≥ 10. We then used juicebox_tools_7.0 hiccups, a functionality of Juicer, to call loops on the .hic files, using default parameters. Juicebox v1.5 was used to visualize the data. We then utilized the InteractionSet library [52] for custom downstream analyses.

Interaction Calling via Fit-Hi-C

To call significant interactions in our 5kb *cis* interaction matrices, we used the Bioconductor version of Fit-Hi-C (https://bioconductor.org/packages/release/bioc/html/FitHiC_.html)[53] with the following parameters to analyze mid-range interactions up to 20Mb: distUpThresh=20000000, noOfBins=200. As input, we used concatenated raw autosomal 5kb *cis* interaction matrices and concatenated autosomal ICE bias vectors generated from our custom ICE implementation. After running Fit-Hi-C, we converted *cis* interactions with *q*-value < 0.01 into InteractionSet[52] format for interrogations of significant interactions using R.

Data Visualization

All visualizations of Hi-C data were generated either using custom R scripts written in house or using Juicebox.[51] For custom visualizations, we used the Bioconductor package ggbio.[54]

RNA-seq Methodology

Total RNA from isolated adult cardiomyocytes was isolated using the Qiagen RNeasy Mini Kit (Cat# 74104). RNA quality was assessed using the Agilent 2100 Bioanalyzer (Cat# G2940CA). Ribosomal RNA was then removed using an Illumina Ribo-Zero rRNA Removal Kit (Cat# MRZH11124) and the RNA-seq library was prepared at the UCLA Clinical Microarray Core using the KAPA Stranded RNA-Seq Library Preparation Kit (Cat# KK8401). The average insert size of each library was measured using an Agilent 4200 TapeStation (Cat# G2991AA). All libraries ranged from 200 to 400bp in length. RNA-seq

libraries were sequenced on two lanes of an Illumina HiSeq 2000 to obtain 2x100bp (paired-end) reads. On average, ~50 million raw pairs were generated per library.

Bioinformatics Analyses of RNA-seq Data

Raw paired-end 100bp reads were sequenced on two lanes of an Illumina HiSeq 2000, and FASTQ files were demultiplexed using a custom Python script. Reads were mapped to the mm10 genome with TopHat v2.1.0,[55] using mm10 bowtie2 indices and a reference GTF provided by iGenomes (Illumina, downloaded from Ensembl release 81 on July 17, 2015). Transcripts were assembled using Cufflinks[56] v2.2.1 with the -G parameter (using the same GTF), and merged using the Cuffmerge feature of Cufflinks with the -g and -s parameters (using the same GTF and the reference FASTA provided by iGenomes, respectively).[56, 57] Cuffquant (from Cufflinks v2.2.1) was performed to determine transcript abundances for all samples, using the -b parameter (with reference FASTA from iGenomes), and the --max-bundle-frags 5000000 and -u parameters. Cuffdiff (also from Cufflinks v2.2.1) was used with identical extra parameters as Cuffquant to determine differential gene expression between conditions and assign *p*-values and *q*-values for each gene. Visualization of RNA-seq data was performed using cummerbund,[58] ggbio,[54] as well as custom R scripts.

ChIP-seq Methodology

ChIP for CTCF (Active Motif Cat# 61311) was performed on pooled isolated cardiomyocytes (~10 million total), using the ChIP-IT High Sensitivity Kit as described.[41] Cardiomyocyte pools were obtained from 3 control and 3 CTCF KO hearts. ChIP-seq

libraries were prepared using the NuGEN Ovation Ultralow System V2 1–16 kit (Cat# 0344).

Bioinformatics Analyses of ChIP-seq Data

ChIP-seq libraries were sequenced on an Illumina HiSeq 2000 instrument. Raw single end 50 bp reads were demultiplexed using a custom Python script and then mapped to the same mm10 Ensembl reference genome as the RNA-seq data, but using Bowtie2[59] v2.2.6 with default parameters. Samtools[60] v0.1.19 was used to convert SAM files to BAM format, and peak calling was performed using the macs2.1.1.2016.0309 callpeak function[61] with default parameters. The CTCF ChIP BAM file was used as the treatment file and input BAM file as the control. A filtering step was added in R to only keep autosomal peaks.

ChIP-seq Visualizations

To visualize the distribution of CTCF peak occupancy across features of the mm10 genome, we used the ChIPseeker[62] package in Bioconductor. We used ngs.plot[63] to generate coverage heatmaps around TSSs and insulation boundaries.

KEGG Analysis

To determine the gene ontology of genes of interest, we used a custom R script that uses KEGG.db in Bioconductor,[64] with custom visualization.

Cardiomyocyte Isolation

Using an established protocol,[65] adult mice were treated with heparin (100 USP units) for 20 minutes to prevent blood coagulation, and then anesthetized with sodium pentobarbital (100μl of 50mg/ml dilution, intra-peritoneal). Upon loss of rear foot reflex, the heart was removed and instantaneously arrested in ice-cold phosphate buffered saline (PBS) and mounted on a modified Langendorff apparatus. After 5 min of perfusion with Tyrode's solution (130mM NaCl, 5.4mM KCl, 1mM MgCl2, 0.6mM Na2HPO4, 10mM glucose, and 10mM HEPES, pH 7.37, oxygenated with 95% (v/v) O_2-5% (v/v) CO_2) at 37°C, the heart was perfused for 15-30 min with 30 ml Tyrode's containing 20mg collagenase type-II and 3mg protease type-XIV and then washed for an additional 10 min with Krebs buffer (KB) (25mM KCl, 10mM KH2PO4, 2mM MgSO4, 20mM glucose, 20mM taurine, 5mM creatine, 100mM potassium glutamate, 10mM aspartic acid, 0.5mM EGTA, 5mM HEPES, pH 7.18) oxygenated with 95% O_2-5% (v/v) CO_2. Cardiomyocytes were dissociated in KB solution, filtered (100μm strainer) and centrifuged 2 min at 1000xg for further usage. This method obtained cells that were ≥95% cardiomyocytes by visual inspection of rod-shaped cell morphology.

Post-mortem histology, Cell Size Quantification and Fibrosis Detection

Whole hearts were rapidly excised from animals, perfused with 0.1 M potassium (K+) solution, fixed using 4% paraformaldehyde and then embedded in paraffin. Hearts were sectioned into 4μm thick sections. Cell size was determined using NIS-Elements Advanced Research v4.0, on heart slices labeled with wheat germ agglutinin (Thermo

Fisher Scientific Cat# W11262). Fibrotic tissue was visualized using the Abcam Picrosirius Red Stain Kit (Cat# ab150681).

CTCF Knockout Mouse Generation

Ctcf-floxed ($Ctcf^{flox}$) mice, in which loxP sites flank exons 3–12, were previously reported.[66] The floxed allele occurs in the C57BL/6 background. To generate cardiac Ctcf cKO mice, we crossed $Ctcf^{flox/flox}$ mice with transgenic mice expressing the α-myosin heavy chain promoter that directs expression of a tamoxifen-inducible Cre recombinase (MerCreMer).[67] Adult $Ctcf^{flox/flox}$ x α-MHC-$MCM^{+/-}$ mice (8 weeks) were administered tamoxifen (Tx) in the chow (0.4mg/g) for 5 weeks plus 1 week on normal chow to deplete CTCF.

Transverse Aortic Constriction and Echocardiographic Measurements

All animal studies were approved by the UCLA Animal Research Committee in compliance with the NIH Guide for the Care and Use of Laboratory Animals. Only adult mice were used in this study. Adult C57BL/6 male and female mice were subjected to transverse aortic constriction surgery to induce pressure overload and cardiac function was measured before, directly after and once every 5 days for the duration of the experiment as described.[19] Animals were sacrificed for experiments based on changes in ejection fraction and heart to body weight, indicating the presence of pathology.

Human Samples

All human samples used in this study were procured in the Ronald Reagan Medical Center at UCLA following patient consent according to IRB#11-001053-AM-00016. mRNA and protein were isolated from LVAD biopsy cores or explanted hearts using conventional approaches.

Electrophoresis and Western Blotting

Cardiomyocytes, neonatal rat ventricular myocytes or human heart tissue were lysed (50mM Tris pH 7.4/10mM ethylenediaminetetraacetic acid [EDTA]/1% Sodium dodecyl sulfate [SDS]/0.1mM phenylmethanesulfonylflouride/protease inhibitor cocktail pellet (Roche)/0.2mM sodium orthovanadate/ 0.1mM sodium fluoride/10mM sodium butyrate), sonicated, and separated via SDS-PAGE using Laemmli buffer. Detection was performed on the LI-COR odyssey. Antibodies were as follows: CTCF mix 1:500 (Abcam, ab70303), (Abcam, ab128909), (BD, 612148), (Abiocode, R3171-1), (Diagenode, C15410210-50) and (Active Motif, 61311), H3K27me3 1:1000 (Abcam, ab6002), H3K4me3 1:1000 (Abcam, ab8580), Rad21 1:1000 (Abcam, ab992), GAPDH 1:1000 (Santa Cruz Biotechnology, sc20357), Actin 1:1000 (Santa Cruz Biotechnology, sc1616), secondaries 1:10,000 (LI-COR, IRDye conjugated). Relative quantification for CTCF expression normalized with respect to Actin was performed using ImageJ software.

RNA Isolation and qPCR

RNA from left ventricular adult cardiomyocytes were isolated using TRIzol (Thermo Fisher Scientific, 15596018) according to the manufacturer's protocol. cDNA was synthesized

using iScript cDNA Synthesis Kit (Bio-Rad, 170-8891). qPCR was performed using SsoFast EvaGreen Supermix (Bio-Rad, 172-5201) on a BioRad, C1000 thermocycler.

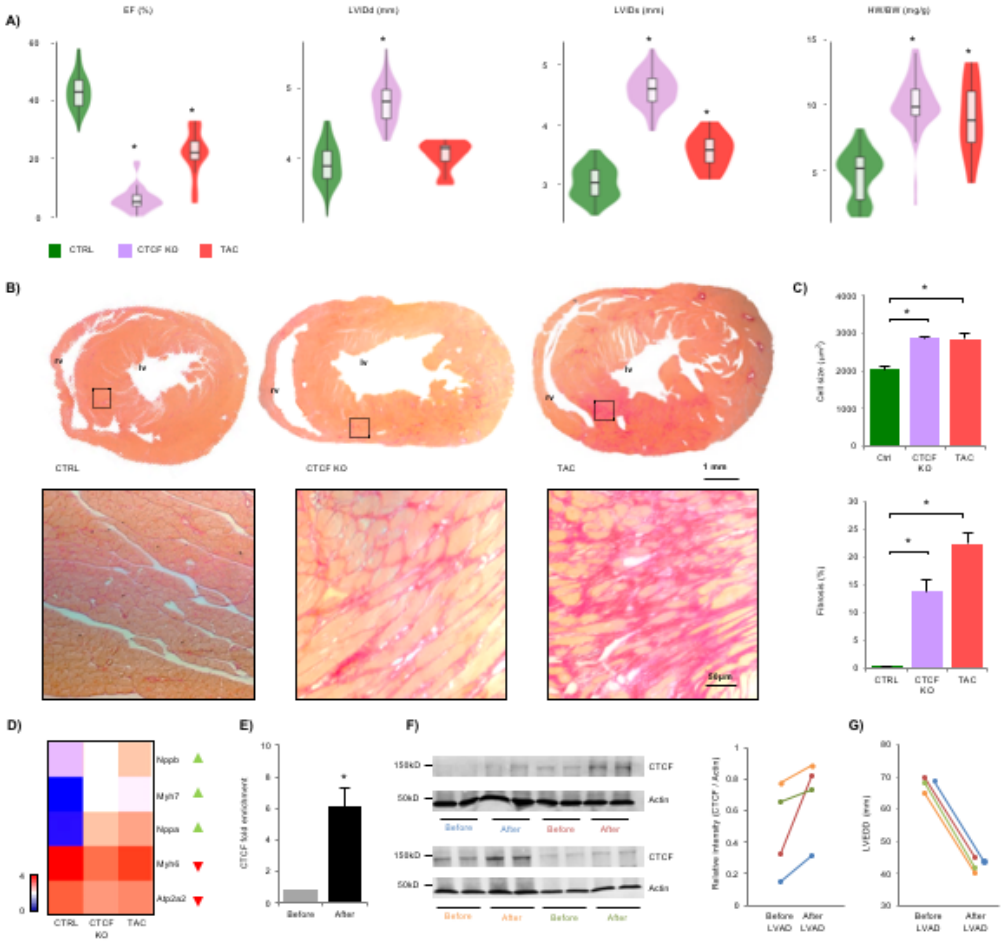


Figure 2-1. Loss of CTCF induces cardiac pathology. (A) Echocardiography measurements of ejection fraction (EF) and LV internal diameter at diastole or systole (LVIDd or LVIDs) demonstrate impairment of EF and chamber dilation in CTCF-KO (purple) and TAC (red) as compared to control (CTRL; green). Heart weight to body

weight ratio (HW/BW) indicates cardiac hypertrophy in CTCF-KO and TAC mice (*n*≥25/group; * p<0.01, Student's *t*-test). **(B)** Picrosirius red staining shows fibrosis after CTCF depletion or TAC (*n*=3/group). **(C)** Top, mean cardiomyocyte area (*n*=20 visual fields of wheat germ agglutinin stained sections across 3 mice per condition); bottom, quantitation of fibrosis from picrosirius red sections (*n*=3/group; * *p*<0.01, Tukey HSD test). **(D)** Stress response gene expression ($\log_{10}$(FPKM+1)). **(E)** Real time qPCR measurements of CTCF levels in human myocardium before and after LVAD (before values normalized to 1 on a per patient basis; n=4, * p<0.01 Student's t-test, bars SD). **(F)** Western blots of CTCF levels in individual patients before and after LVAD (left); quantitation of western blots normalized to actin on a per patient, per sample basis (right). **(G)** Left ventricular end diastolic dimension measurements before and after LVAD. Color-coding in E-G indicates separate patients. Lines shifted horizontally in G for ease of viewing.

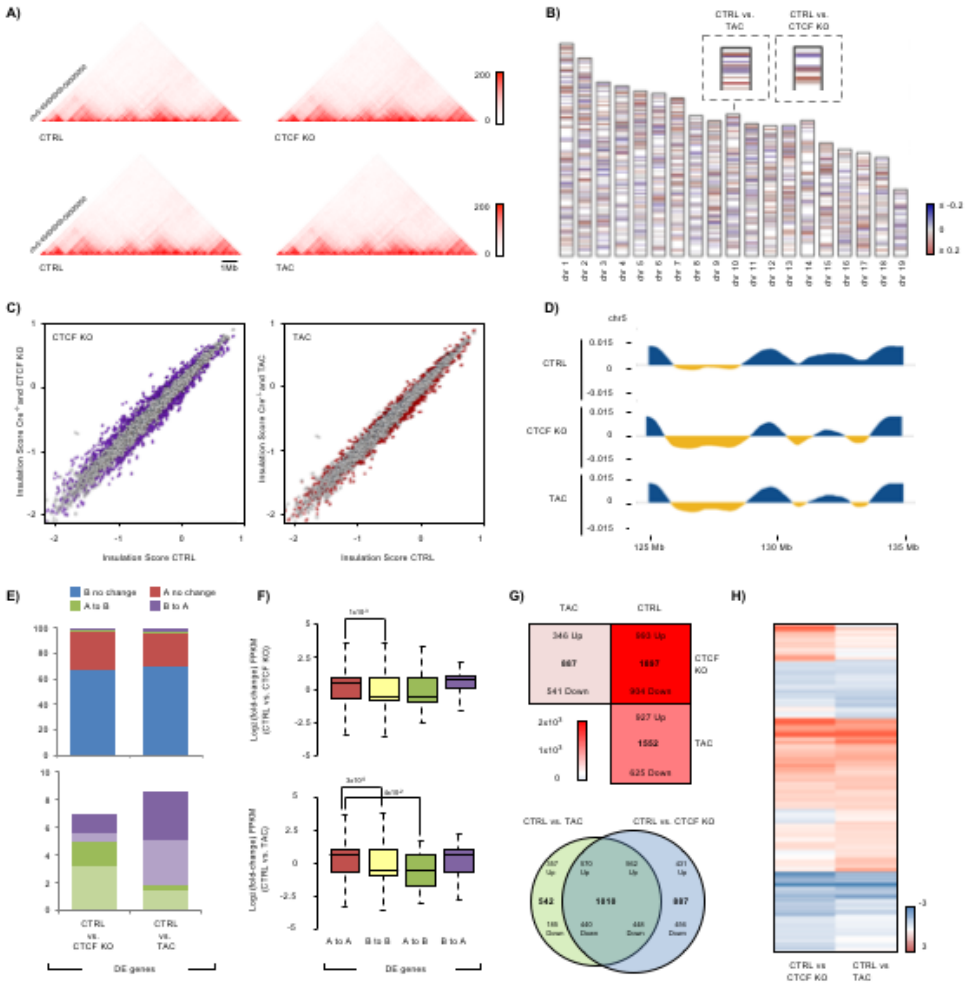


Figure 2-2. High-resolution cardiac chromatin conformation analyses reveal changes in chromatin compartmentalization and gene expression in pressure overload and CTCF-KO mice. (A) Structure of topological associating domains (TADs) is revealed from contact frequency heatmaps showing *cis* interaction profile on example

chromosome 5 for control, TAC and CTCF-KO chromatin. **(B)** Strength of boundaries between TADs are displayed for all chromosomes comparing control and TAC (red, higher; blue, lower; Insets: example region on chromosome 10 in control vs. TAC and control vs. CTCF-KO; Figure 2-13A shows control vs. CTCF-KO for all chromosomes). **(C)** Quantitation of insulation score differences in control versus KO (left) or TAC (right). Colored dots indicate significant changes (grey dots show range of variation between two control conditions: untreated wild type mouse and untreated $Cre^{+/-}$ mouse). **(D)** A/B compartmentalization, an indicator of genome accessibility at individual loci, for an example region on chromosome 5 is plotted in blue (open, A) and yellow (closed, B): CTRL A/B status on top, CTCF-KO in middle, and TAC on bottom. **(E)** Quantification of the genome-wide changes in A/B compartment change with CTCF depletion (left) or TAC (right). Bottom panels highlight only bins that change compartment; dark and light colors represent up- or down-regulated genes, respectively. **(F)** Relationship of compartmentalization to gene expression is measured in TAC or CTCF-KO hearts. Log_2(fold-change) of FPKM for the differentially expressed genes that either remain in the same compartment or that change compartments with CTCF depletion (top) and TAC (bottom). *p*-values via Wilcoxon test; whiskers indicate interquartile range. **(G)** Heat matrix (top) showing number of differentially expressed genes between CTRL, CTCF KO, and TAC (intensity indicates number of genes). Venn diagram (bottom) showing overlap of differentially expressed genes between CTRL vs. CTCF KO and CTRL vs. TAC. **(H)** Heatmap depicting log_2(KO/CTRL) and log_2(TAC/CTRL) FPKM for the differentially expressed genes with $q < 0.01$ (left). (FPKM) fragments per kb of exon per million mapped reads.

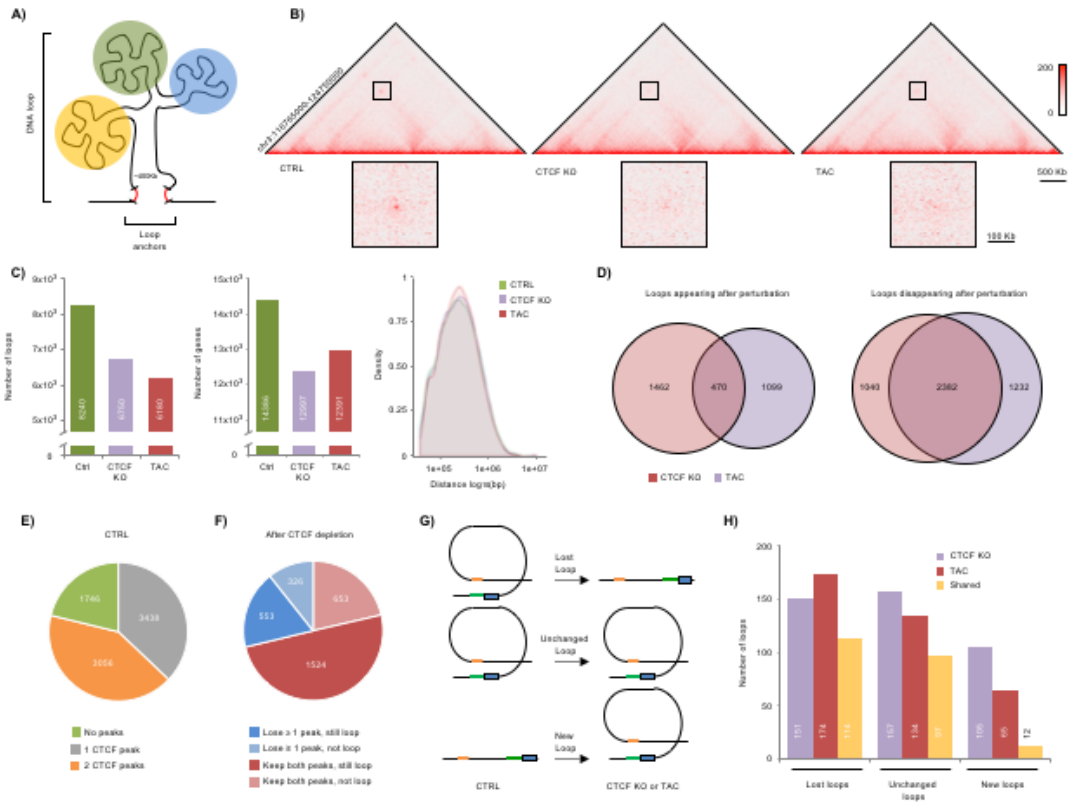


Figure 2-3. Short- and long-range chromatin interactions, and stable chromatin looping, are altered after pressure overload or CTCF-KO. (A) Schematic displaying that loops are demarcated by two anchors and contain regions of high-frequency interactions, indicated by the colored circles. **(B)** The bioinformatic tool Juicebox is used to display an example loop that is lost with CTCF depletion (middle) and TAC (right). **(C)** Quantitation of the phenomenon in (B) across genome, showing number of chromatin loops (left), number of genes within loops (middle), and loop sizes (right; CTRL, green; CTCF-KO, purple; TAC, red). **(D)** Overlap of loops only appearing in CTCF-KO or TAC

when compared to CTRL (left); overlap of loops that disappear in CTCF-KO and TAC (right). **(E)** CTRL loops in which zero (green), one (grey), or both (yellow) anchors overlap with a CTCF peak. **(F)** CTRL loops that lose ≥ 1 CTCF peak during CTCF depletion (blue) and CTRL loops that keep both CTCF peaks during CTCF depletion (red). Darker shade indicates loops that were preserved with CTCF depletion, while lighter color indicates loops that were lost. **(G)** Schematic demonstrating types of alterations in looping architecture that can occur, including loss of loops mediating enhancer-promoter interactions (top), no change (middle), or formation (bottom). Enhancers are orange, promoters green and genes blue. **(H)** Quantification of changes in enhancer-promoter loops.

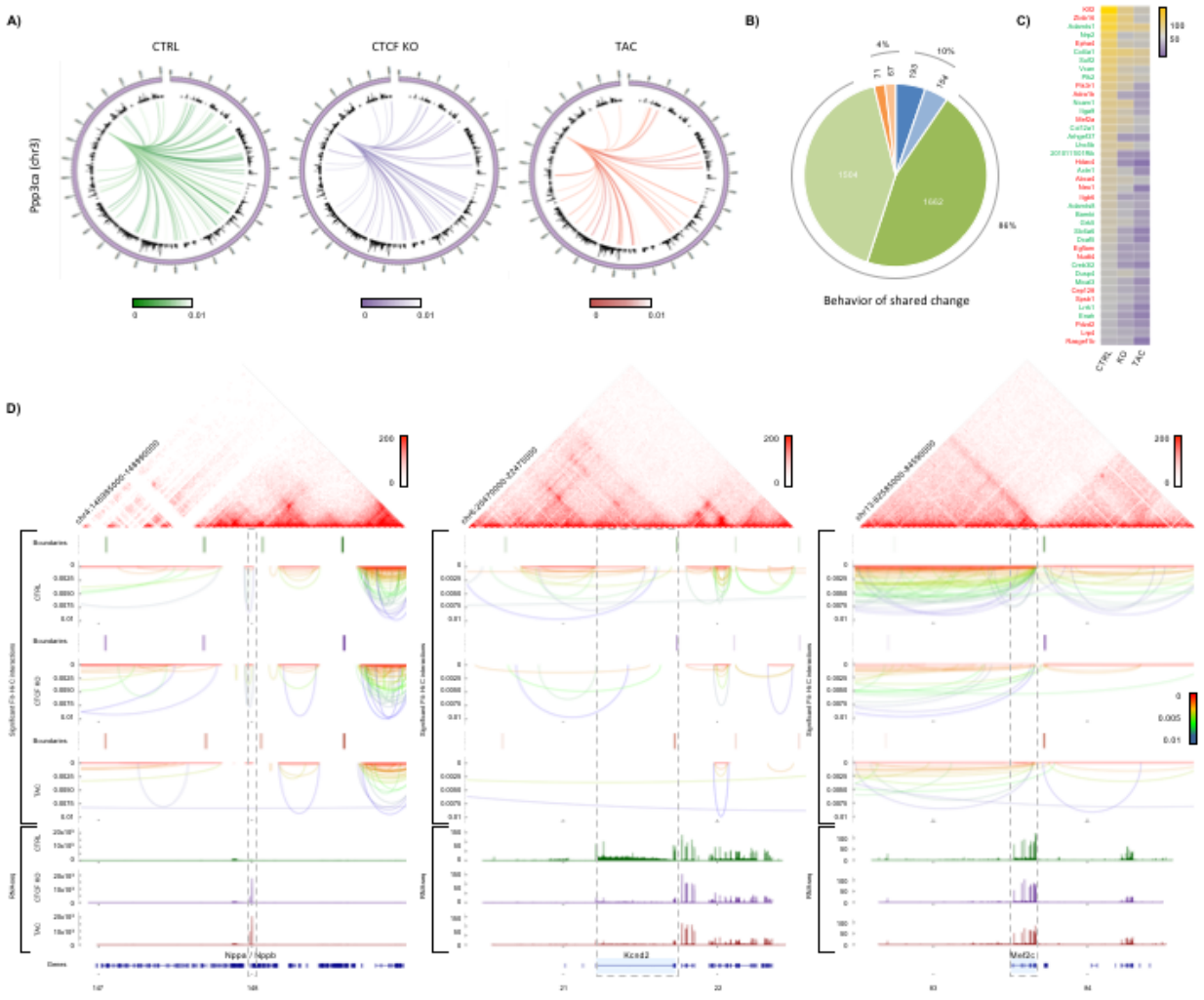


Figure 2-4. Chromatin architecture is remodeled around cardiac genes during disease. (A) As an example cardiac disease gene with changing long range, intra-chromosomal contacts, the interactions emanating from *Ppp3ca* in control and after CTCF depletion or TAC are shown (q<0.01; 40kb resolution; outer circle, chromosome position;

black, mm10 genes). **(B)** Examining gene expression data for all 5443 genes with shared interaction behavior between CTCF-KO and TAC, 3651 genes were found to be differentially expressed in the same direction in perturbations compared to control. Most (86%) of these gene expression changes were associated with decreased chromatin interactions (green colors), with 1662 up regulated (dark green) and 1504 down regulated (light green). The remainder of the interaction changes (14%) were distributed between the other possible scenarios: increased interactions and expression (193, dark blue), increased interaction and decreased expression (154, light blue), no change in interactions and either an increase (71, dark orange) or decrease (67, light orange) in expression. **(C)** Heatmap showing the number of significant (q<0.01) interactions overlapping with differentially expressed genes; top 40 shared differentially expressed genes shown, sorted by number of significant interactions in CTRL (gene name labeling shows direction of expression, where green is up and red is down). **(D)** Higher resolution mapping of local neighborhood interactions for the example genes *Nppa*/*Nppb*, *Kcnd2*, and *Mef2c* gene loci +/- 1Mb (q<0.01). Lines revealing precise contact sites are color-coded by q-value significance, with red being the most significant. RNA-seq tracks depict gene expression for CTRL (green), CTCF-KO (purple), and TAC (red).

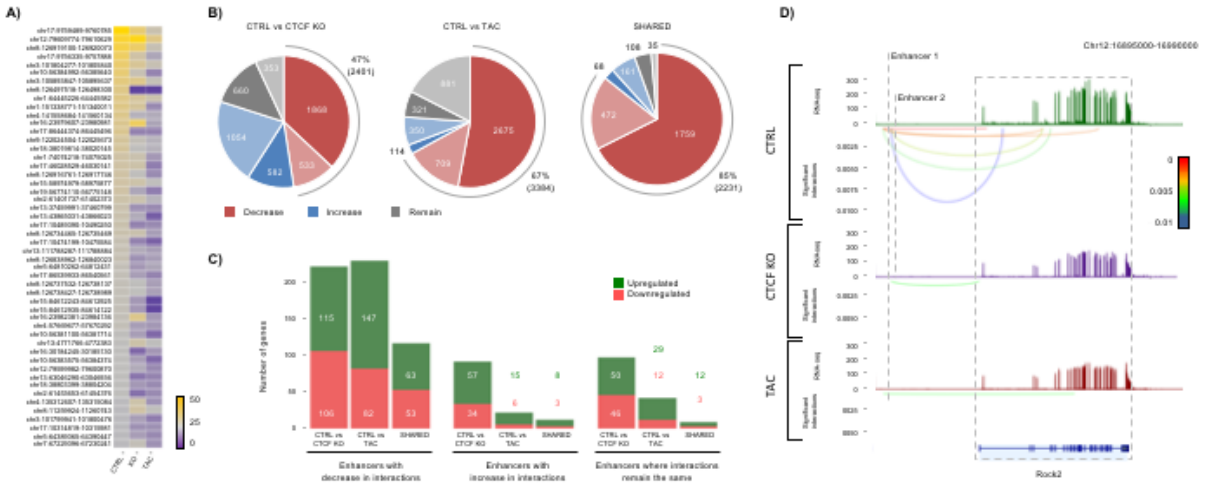


Figure 2-5. Three-dimensional interactions between enhancers and genes are restructured after pressure overload or CTCF-KO. (A) Number of contact sites per cardiac enhancer are quantified (q<0.01; rows are top 50 enhancers sorted by number of interactions). **(B)** Top, enhancers in which the number of significant (q<0.01) interactions (determined by the Fit-Hi-C tool) increases (blue), decreases (red), or remains the same (grey) with CTCF-KO (left) TAC (center) and with consistent changes with CTCF-KO or TAC (right). Darker shading indicates enhancers that interact with genes; lighter colors

are regions not annotated as coding. **(C)** Number of genes that interact with enhancers from (B), stratified by whether their expression is upregulated (green) or downregulated (red). Groupings are separated into enhancers whose number of overlapping interactions decreases (left), increases (middle) or remains the same (right) after perturbation. **(D)** Contact site mapping of interactions is shown for the example gene *Rock2*, which exhibits decreased interaction with enhancers in CTCF-KO and TAC, concomitant with decreased expression.

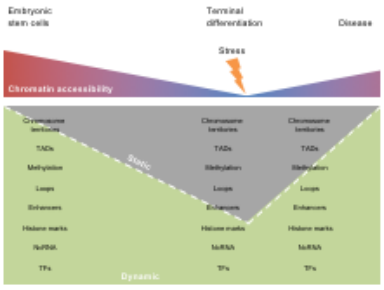
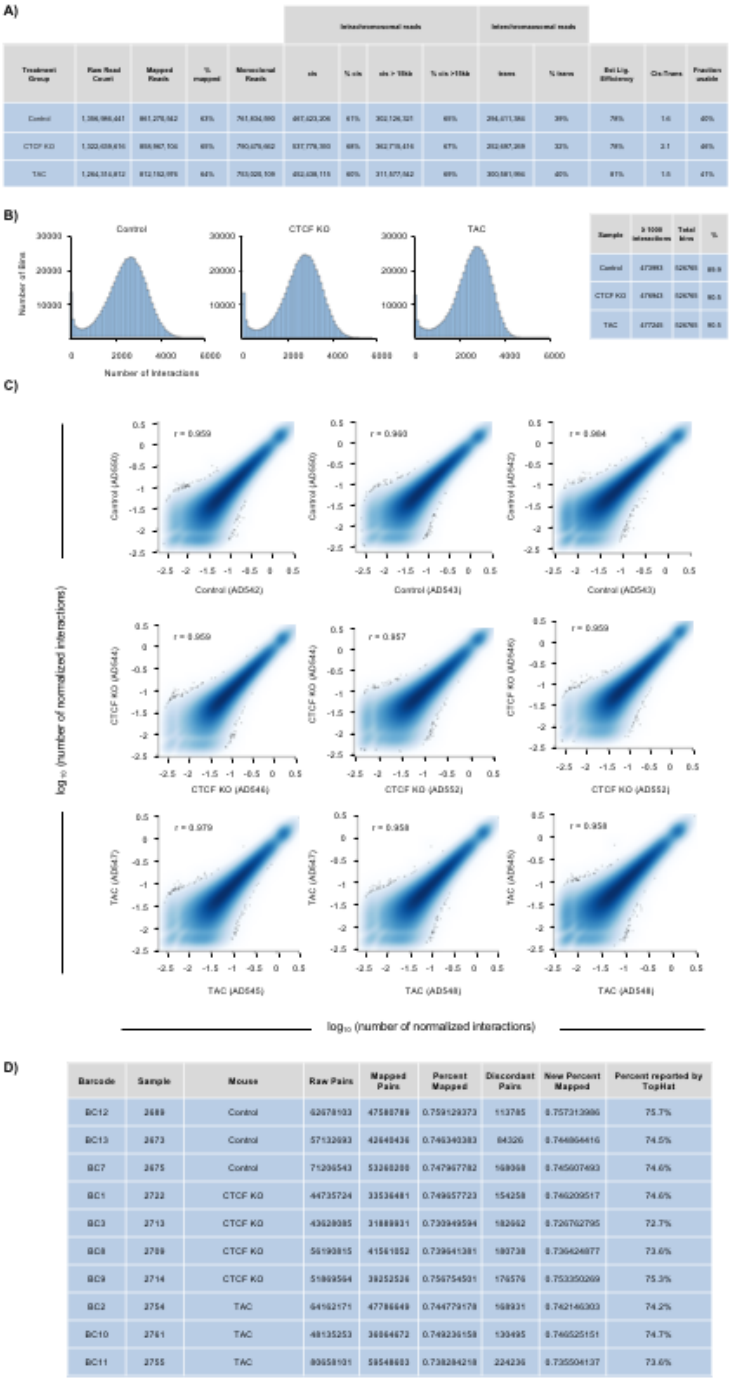


Figure 2-6. Model for epigenomic changes in development and disease. Development is accompanied by changes in chromatin to endow terminally differentiated cells with stable transcriptomes. Disease upsets this balance, transitioning select regions of the genome into more dynamic conformations through effects on chromatin structure, enhancer-gene looping, histone modifications, DNA methylation and other factors. This model is based on findings from this paper and previous publications[2-5, 10, 11, 15-17, 29-33] as described in text.



Barcode	Sample	Mouse	Raw Pairs	Mapped Pairs	Percent Mapped	Discordant Pairs	New Percent Mapped	Percent reported by TopHat
BC12	3689	Control	62678103	47580789	0.759129373	113785	0.757313986	75.7%
BC13	3673	Control	57132993	42648436	0.746349383	84326	0.744864416	74.5%
BC7	3675	Control	71206543	53268288	0.747967782	169968	0.745607493	74.6%
BC1	3722	CTCF KO	44735724	33536481	0.749557723	154258	0.746209517	74.6%
BC3	3713	CTCF KO	43628085	31888931	0.730949594	182662	0.726762795	72.7%
BC8	3709	CTCF KO	56190815	41561852	0.739941381	193738	0.736424877	73.6%
BC9	3714	CTCF KO	51999564	39252526	0.756754501	176576	0.753350269	75.3%
BC2	3754	TAC	64162171	47786649	0.744779178	169931	0.742146303	74.2%
BC10	3761	TAC	48135253	36064672	0.749236158	130495	0.746525151	74.7%
BC11	3755	TAC	80658101	59548883	0.738284218	224236	0.735504137	73.6%

Figure 2-7. Genome-wide chromatin conformation capture datasets and quality control analysis for Hi-C and RNA-seq. **(A)** Table summarizing read counts and mapping statistics for each sample type. Replicates of the same condition were combined

upon verification that interactions and AB compartmentalization correlated with Pearson correlation > 0.90 (at 5kb resolution). **(B)** Coverage histograms of raw 5kb *cis* and *trans* matrices for each sample type (x axis indicates number of raw interactions and y axis indicates the number of bins that had such number of interactions). **(C)** Normalized interaction frequency per bin is highly conserved between biological replicates, Pearson correlation coefficient of > 0.95 in all comparisons. For this analysis, cis interactions ≤ 2Mb were compared. The normalized *cis* matrices were generated from ice-normalized *cis* and *trans* genome-wide matrices that were divided by matrix sum and multiplied by the arbitrary value of 1 million. **(D)** Summary statistics of RNA-seq sequencing data, including number of mapped read pairs for each sample.

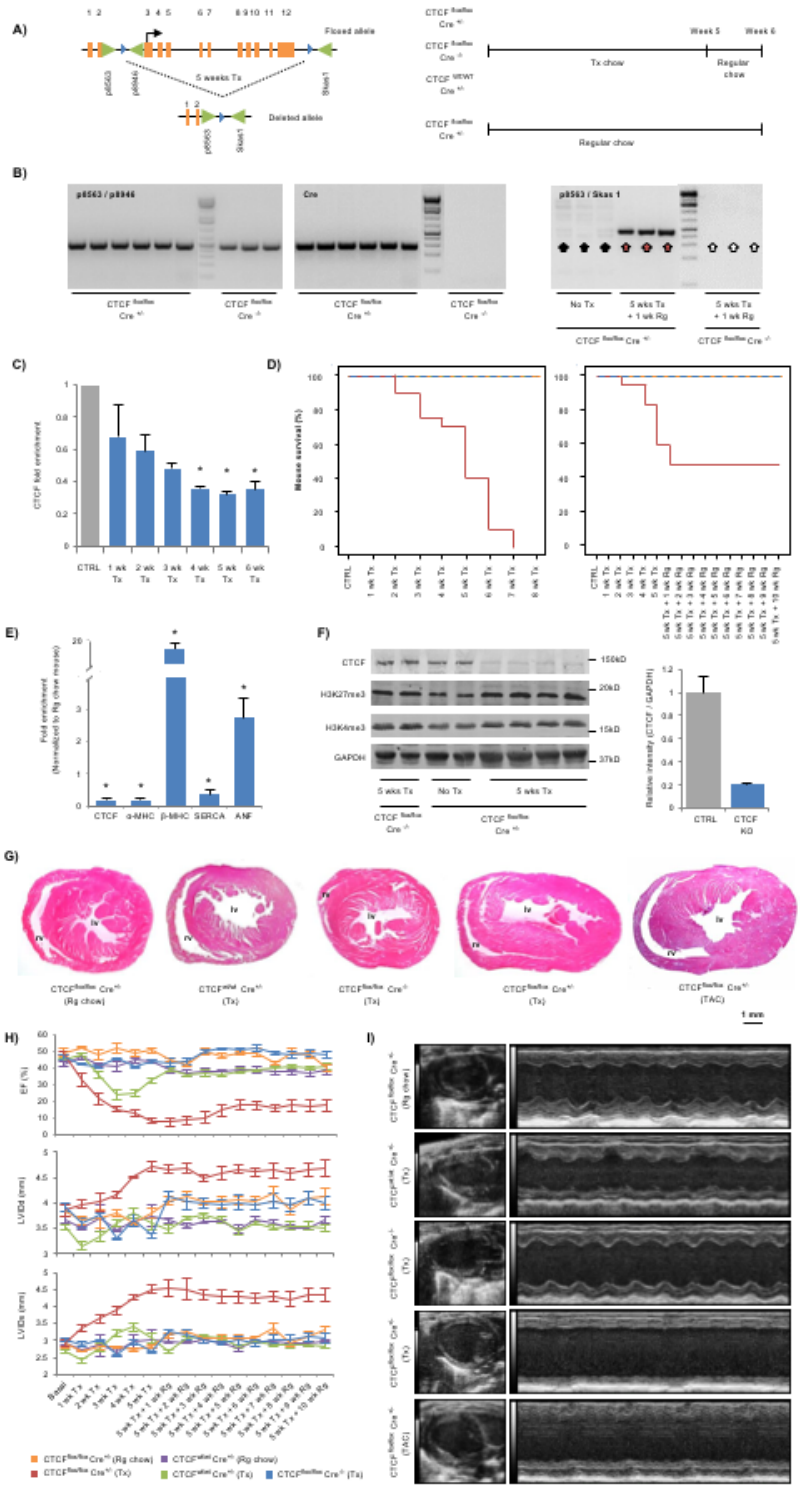


Figure 2-8. Generation and phenotypic characterization of cardiac-specific CTCF knockout mice. (A) Left, diagram showing the floxed *ctcf* gene from exon 3 to 12.

Labeled yellow bars mark exons 1 to 12 and labeled green arrows indicate primers used to verify CTCF depletion via PCR. Primers p8563 and Skas1 should have a PCR product when exons 3 to 12 are excised by the Cre recombinase via the LoxP sites (blue triangles), and Primers p8563 and p8946 should have a PCR product in any scenario in which the Cre recombinase is not expressed. Right, α-MHC-$Cre^{+/-}$ x $CTCF^{flox/flox}$ mice were fed Tx containing chow (0.4mg Tx/g chow) for 5 weeks and then switched to normal chow for 1 week, resulting in deletion of exons 3-12 of CTCF. α-MHC-$Cre^{+/-}$ x $CTCF^{flox/flox}$ mice fed normal chow for 6 weeks, and α-MHC-$Cre^{+/-}$ x $CTCF^{WT/WT}$ mice fed tamoxifen chow for 5 weeks plus 1 week on normal chow, served as controls. **(B)** PCR genotyping confirms the presence of the floxed CTCF alleles in our $CTCF^{flox/flox}$ $Cre^{+/-}$ mice (left), as well as presence of the *Cre* gene in these mice (center). After 5 weeks of tamoxifen treatment followed by 1 week on regular chow, gene excision in the heart was confirmed by PCR using p8563 and Skas1 primers (right). PCR amplified bands at ~500bp in the knockout (red arrows), but did not amplify the ~27kb region of the floxed allele in $Cre^{-/-}$ control mice (white arrows) or normal chow fed mice (black arrows). (Rg) regular chow. **(C)** CTCF KO was confirmed by RT-qPCR (left panel) for α-MHC-$MCM^{+/-}$ x $CTCF^{flox/flox}$ mice that were fed tamoxifen for 1, 2, 3, 4, 5 and 6 weeks. Blue bars show fold enrichment when compared to CTRL, and error bars indicate standard deviation; * indicates $p<0.05$ Student's t-test. **(D)** Kaplan-Meier survival curves of $CTCF^{flox/flox}$ $Cre^{+/-}$ (n=20) (red line), $CTCF^{wt/wt}$ $Cre^{+/-}$ (n=10) (green line), $CTCF^{flox/flox}$ $Cre^{-/-}$ (n=10) (blue line) fed with Tx and $CTCF^{flox/flox}$ $Cre^{+/-}$ (n=10) (orange line) fed with regular chow. Left panel, 5 weeks after Tx treatment, mouse survival decreased 63%, while survival dropped to 8% at 6 weeks. Right panel, Mouse survival stabilizes at 53% after the initial 5 weeks when mice are

returned to normal chow. (Rg) regular chow. **(E)** RT-qPCR revealed mRNA down-regulation of SERCA (*Atp2a2*) and α-MHC (*Mhy6*), and upregulation of ANF (*Nppa*) and β-MHC (*Mhy7*), in mice undergoing 5 weeks Tx treatment plus 1 week regular chow (Rg). Bars show fold enrichment when compared to CTRL mouse given regular chow, and error bars indicate standard deviation; * indicates $p<0.05$ Student's t-test. **(F)** Western blot confirming CTCF depletion in tamoxifen-treated $CTCF^{flox/flox}$ $Cre^{+/-}$ mice (left), and showing no change in H3K27me3 and H3K4me3 with CTCF depletion. Quantification of average relative CTCF intensity shows a decrease in CTCF protein levels in $CTCF^{flox/flox}$ $Cre^{+/-}$ mice after 5 weeks of tamoxifen treatment. **(G)** Representative hematoxylin and eosin staining, comparing $CTCF^{wt/wt}$ $Cre^{+/-}$, $CTCF^{flox/flox}$ $Cre^{-/-}$and $CTCF^{flox/flox}$ $Cre^{+/-}$ after 5 weeks Tx treatment followed by 1 week on regular chow. $CTCF^{flox/flox}$ $Cre^{+/-}$ treated with regular chow for 6 weeks were used as an extra negative control while TAC samples were used as a positive control for disease. Considerable dilatation of the ventricles is observed in the tamoxifen-treated $CTCF^{flox/flox}$ $Cre^{+/-}$ mice as well as the TAC mice. *n* = 5 for each condition. (lv) left ventricle, (rv) right ventricle, (Rg) regular chow. Scale bar: 1mm. **(H)** Echocardiographic measurements of all the mice used in the study. Notably, $CTCF^{flox/flox}$ $Cre^{+/-}$ mice treated with tamoxifen had a pathological phenotype when compared to the $CTCF^{flox/flox}$ $Cre^{+/-}$ or $CTCF^{wt/wt}$ $Cre^{+/-}$ that were given regular chow as well as the $CTCF^{wt/wt}$ $Cre^{+/-}$ or $CTCF^{flox/flox}$ $Cre^{-/-}$ mice treated with tamoxifen. *n* = 8-17 depending on the survival rate for each condition. Line indicates mean value for the parameter of interest while the error bars indicate standard deviation. (EF) ejection fraction, (LVIDd) LV internal diameter at diastole, (LVIDs) LV internal diameter at systole. **(I)** Representative B-mode (left) and M-mode (right) echocardiography images at 6 weeks from normal chow $CTCF^{flox/flox}$

Cre+/- (upper panel) or tamoxifen fed $CTCF^{wt/wt}$ $Cre^{+/-}$, $CTCF^{flox/flox}$ $Cre^{-/-}$and $CTCF^{flox/flox}$ $Cre^{+/-}$ mice. $CTCF^{flox/flox}$ $Cre^{+/-}$ mice show pathological phenotype as in (H); a TAC mouse is shown for comparison.

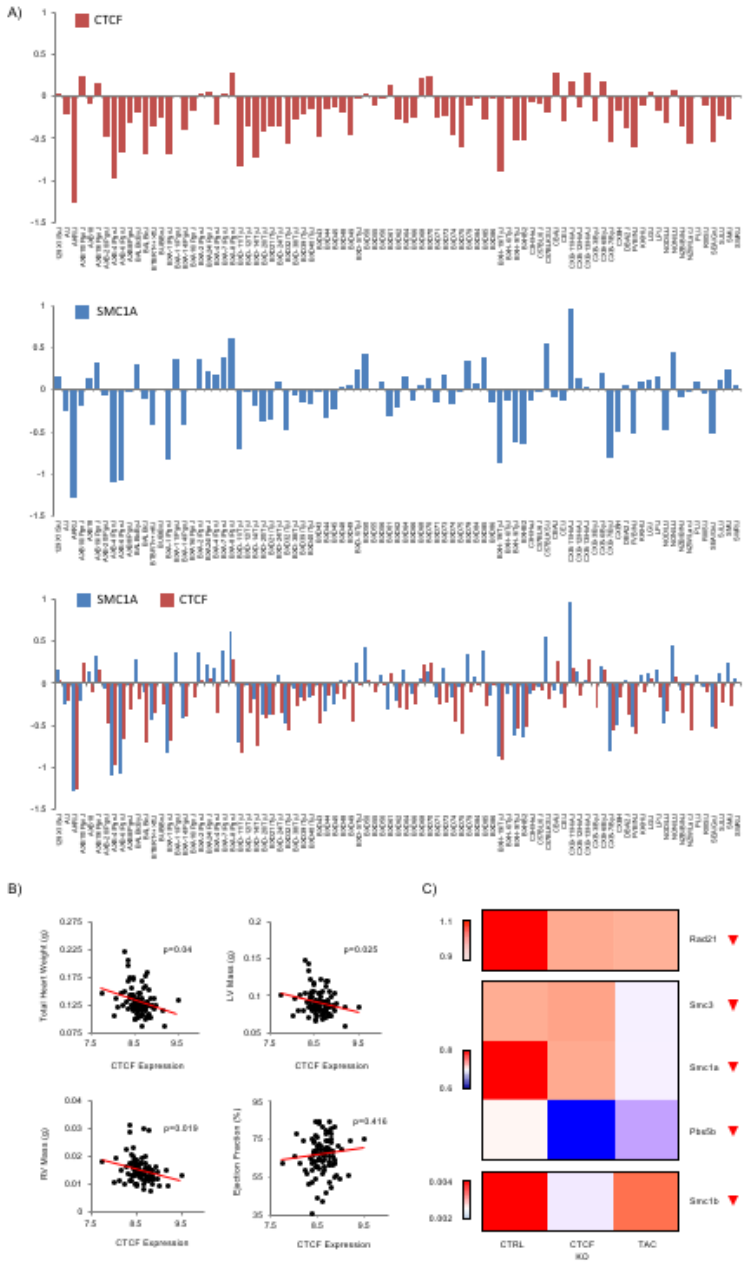


Figure 2-9. CTCF expression is down-regulated by pathologic stress and inversely correlated with heart size. (A) CTCF levels in mice undergoing a model of isoproterenol-

induced heart failure across 87 genetically distinct strains. CTCF mRNA levels decrease across 80% of the mouse strains after isoproterenol treatment (top and middle panels; data from ref. 68). Cohesin subunit Smca1 shows coordinate regulation with CTCF in this model. **(B)** Phenotype data from 87 mouse strains treated with ISO are plotted against CTCF mRNA expression after ISO. Total heart mass, left ventricular mass, and right ventricular mass all show significant negative correlation with CTCF expression, while ejection fraction shows no correlation. p-values were calculated using the Pearson correlation coefficient. (LV) left ventricle, (RV) right ventricle (data from ref. 68). **(C)** Expression levels of different cohesin subunits decrease after CTCF depletion and TAC (log_{10}(FPKM+1).

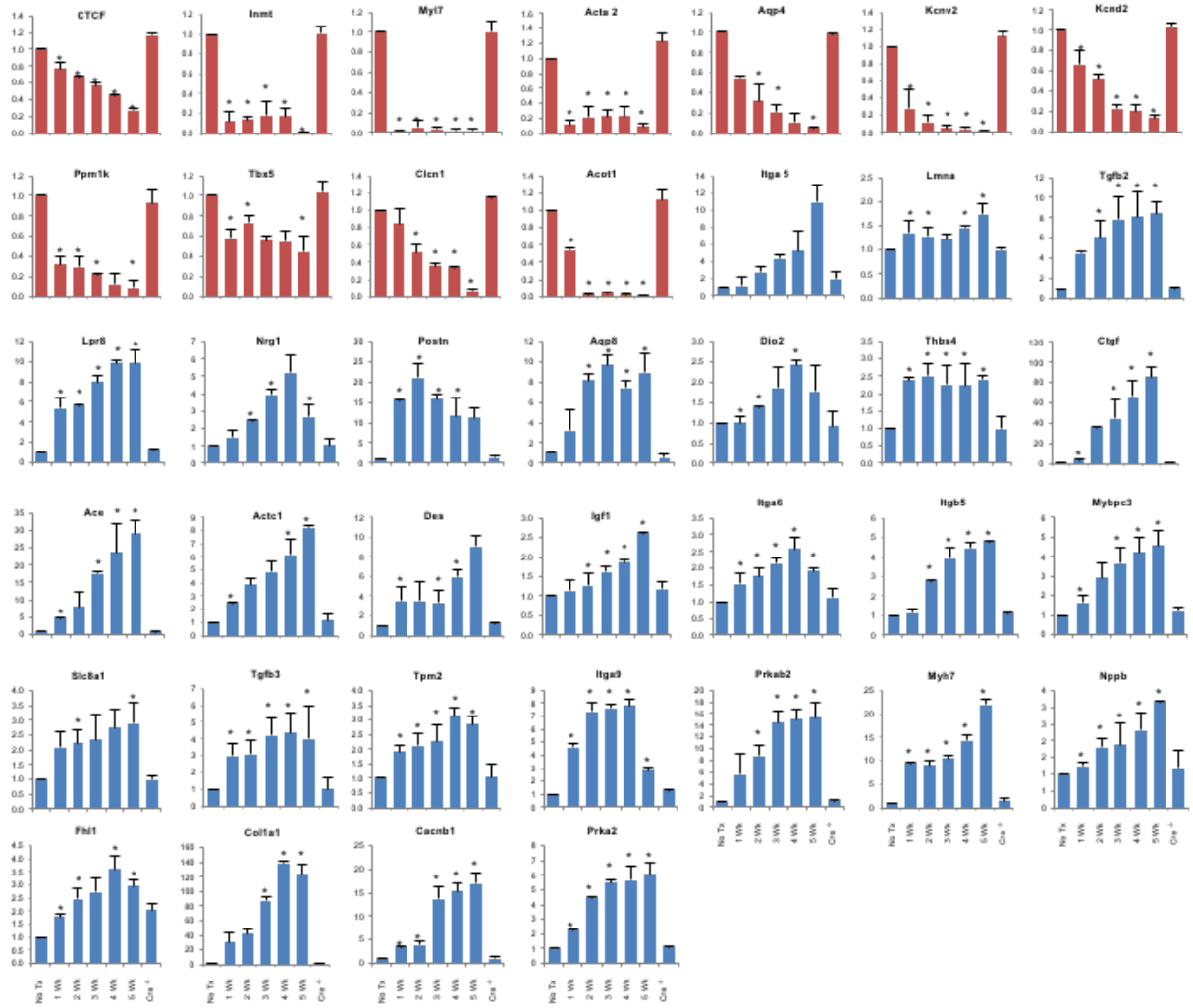


Figure 2-10. Cardiac disease associated genes differentially expressed after CTCF depletion. Real time qPCR measuring expression levels of different cardiac genes after 1, 2, 3, 4 and 5 weeks Tx treatment. Mice fed with regular chow (No Tx) and those fed

with Tx were used as controls. Red and blue coloring indicates down- or up-regulation, respectively. (n=4/group; * $p<0.05$, Student's t-test).

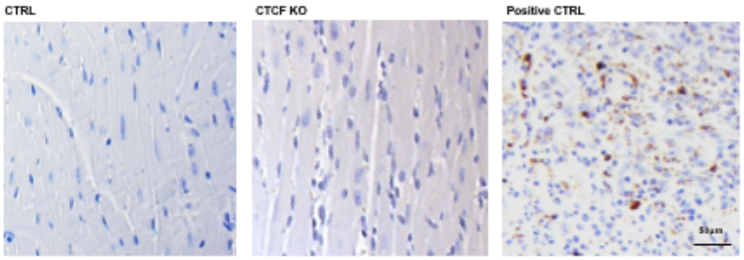


Figure 2-11. CTCF depletion does not promote apoptosis. Cleaved caspase-3 staining of heart sections from CTRL and CTCF KO mice. Human sarcoma tumor cells were used as a positive control (Right panel).

Sample	Management	Sex	Age (Year)	Etiology	EF	Wb quantification (After / Before)	LVEDD (mm)	NYHA class	BNP (pg/mL)	Comorbidities	Medical History	Medications
114	Before LVAD	Male	43	NICM	25	2.12	69	4	402	None noted	Dyslipidemia Atrial fibrillation	Pravastatin Cardizem
119	After LVAD (7 weeks)						45	2	73			
167	Before LVAD	Male	55	NICM	20	2.54	70	4	778	COPD due to 50 years of smoking history	Adenoidectomy Non sustained ventricle tachycardia Heparin induced thrombocytopenia	Amiodarone Epinephrine Alteplase
177	After LVAD (8 weeks)						45	2	59			
141	Before LVAD	Male	65	ICM	20	1.14	65	4	548	STEMI Type 2 DM	Arthritis Hypertension Coronary artery disease	Amlodipine Atorvastatin Hydralazine Metoprolol Sildenafil Spironolactone Warfarin
161	After LVAD (24 weeks)						40	2	57			
117	Before LVAD	Male	55	ICM	20	1.12	68	3B	402	Type 2 DM Renal insufficiency Dyslipidemia	Hypertension Acute Pancreatitis Colitis	Amlodipine Furosemide Pravastatin
181	After LVAD (51 weeks)						42	2	66			

Figure 2-12. Human subject clinical data. Available data on patients from whom heart samples were obtained is provided. “Wb quantitation” is the CTCF protein level measured by Western blot expressed as a ratio of after to before LVAD. NICM, non-ischemic cardiomyopathy; ICM, ischemic cardiomyopathy; NYHA, New York Heart Association Classification; EF, ejection fraction at time of LVAD placement.

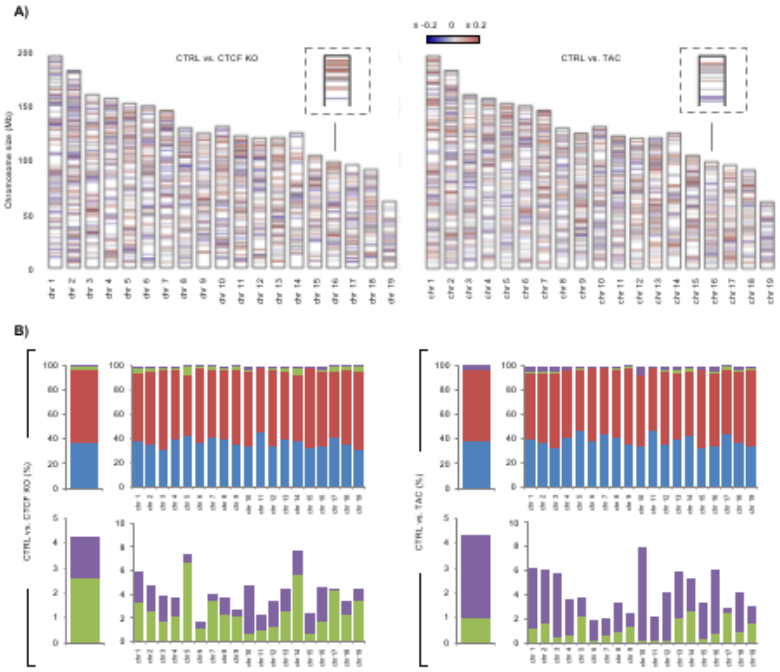


Figure 2-13. Differences in boundary strength and A/B compartmentalization per chromosome. (A) Boundary strength differences between groups (red, higher; blue, lower; Insets: example region on chromosome 16. **(B)** Percent of genome-wide A/B compartment change with CTCF depletion (left) and TAC (right). Bottom panels highlight only bins that change compartment.

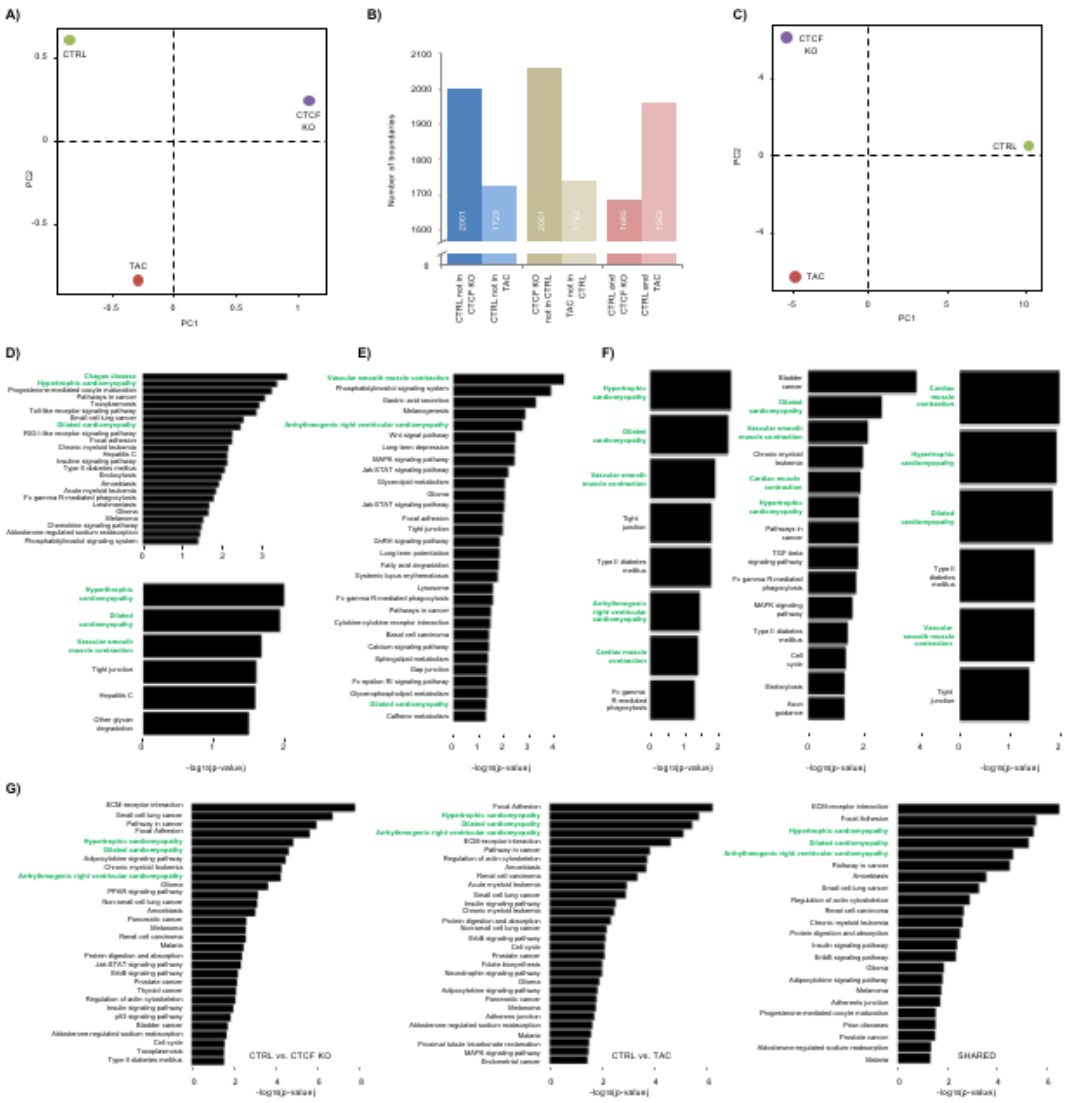


Figure 2-14. Implications for cardiac phenotype from chromatin structural changes in TAC and CTCF-KO cardiac myocytes. (A) Principal component analysis of the shared insulation boundary strengths. **(B)** Number of new boundaries between different

experimental groups. Analyses do not distinguish between boundaries that emerge either from shifting of an existing boundary outside the analysis window (+/-5kb from CTRL boundary center) or *de novo* formation of a completely new boundary. Color codes: Boundaries only found in CTRL (blue), only found in perturbation (beige), and shared between CTRL and perturbation (pink). **(C)** Principal component analysis of the RNA-seq samples. Control separates from CTCF KO and TAC on PC1. **(D)** KEGG pathway analysis of up-regulated (top) or down-regulated (bottom) genes bound by CTCF. Terms highlighted in green are cardiac-related. **(E)** KEGG pathway analysis of genes (2570; Figure 2-3F) that lie within the loops that lose ≥1 CTCF peak in CTCF-KO. **(F)** KEGG pathway analysis of genes (Figure 2-3H, "lost loops") that lie within the 151 enhancer-promoter loops that are lost in CTCF-KO (left), the 174 lost with TAC (middle), and the 114 loops lost in both diseased conditions (CTCF-KO and TAC; right). **(G)** KEGG pathway analysis of the differentially expressed genes (Figure 2-5C, "enhancers with decrease in interactions") interacting with enhancers that lose interactions after perturbation (left, control vs. CTCF-KO; middle, control vs. TAC; right, shared).

Chapter 2: Bibliography

1. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, de Ferranti SD, Floyd J, Fornage M, Gillespie C, Isasi CR, Jimenez MC, Jordan LC, Judd SE, Lackland D, Lichtman JH, Lisabeth L, Liu S, Longenecker CT, Mackey RH, Matsushita K, Mozaffarian D, Mussolino ME, Nasir K, Neumar RW, Palaniappan L, Pandey DK, Thiagarajan RR, Reeves MJ, Ritchey M, Rodriguez CJ, Roth GA, Rosamond WD, Sasson C, Towfighi A, Tsao CW, Turner MB, Virani SS, Voeks JH, Willey JZ, Wilkins JT, Wu JH, Alger HM, Wong SS, Muntner P, American Heart Association Statistics C and Stroke Statistics S. Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association. *Circulation*. 2017;135:e146-e603.

2. Rajabi M, Kassiotis C, Razeghi P and Taegtmeyer H. Return to the fetal gene program protects the stressed heart: a strong hypothesis. *Heart Fail Rev*. 2007;12:331-43.

3. McKinsey TA and Olson EN. Dual roles of histone deacetylases in the control of cardiac growth. *Novartis Found Symp*. 2004;259:132-41; discussion 141-5, 163-9.

4. Molkentin JD and Dorn GW, 2nd. Cytoplasmic signaling pathways that regulate cardiac hypertrophy. *Annu Rev Physiol*. 2001;63:391-426.

5. Kim SY, Morales CR, Gillette TG and Hill JA. Epigenetic regulation in heart failure. *Curr Opin Cardiol*. 2016;31:255-65.

6. Orphanides G and Reinberg D. A unified theory of gene expression. *Cell*. 2002;108:439-51.

7. Strahl BD and Allis CD. The language of covalent histone modifications. *Nature*. 2000;403:41-5.

8. Dekker J and Mirny L. The 3D Genome as Moderator of Chromosomal Communication. *Cell*. 2016;164:1110-21.

9. Heinz S, Romanoski CE, Benner C and Glass CK. The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol*. 2015;16:144-54.

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

11. Papait R, Cattaneo P, Kunderfranco P, Greco C, Carullo P, Guffanti A, Vigano V, Stirparo GG, Latronico MV, Hasenfuss G, Chen J and Condorelli G. Genome-wide analysis of histone marks identifying an epigenetic signature of promoters and enhancers underlying cardiac hypertrophy. *Proceedings of the National Academy of Sciences of the United States of America*. 2013;110:20164-9.

12. May D, Blow MJ, Kaplan T, McCulley DJ, Jensen BC, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Afzal V, Simpson PC, Rubin EM, Black BL, Bristow J, Pennacchio LA and Visel A. Large-scale discovery of enhancers from human heart tissue. *Nature genetics*. 2012;44:89-93.

13. Phillips JE and Corces VG. CTCF: master weaver of the genome. *Cell*. 2009;137:1194-211.

14. Schmitt AD, Hu M, Jung I, Xu Z, Qiu Y, Tan CL, Li Y, Lin S, Lin Y, Barr CL and Ren B. A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell Rep*. 2016;17:2042-2059.

15. 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 and Bruneau BG. Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage. *Cell*. 2012;151:206-20.

16. 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 and Murry CE. A temporal chromatin signature in human embryonic stem cells identifies regulators of cardiac development. *Cell*. 2012;151:221-32.

17. Haldar SM and McKinsey TA. BET-ting on chromatin-based therapeutics for heart failure. *J Mol Cell Cardiol*. 2014;74:98-102.

18. Dixon JR, Gorkin DU and Ren B. Chromatin Domains: The Unit of Chromosome Organization. *Mol Cell*. 2016;62:668-80.

19. Franklin S, Chen H, Mitchell-Jordan S, Ren S, Wang Y and 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.

20. Crane E, Bian Q, McCord RP, Lajoie BR, Wheeler BS, Ralston EJ, Uzawa S, Dekker J and Meyer BJ. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature*. 2015;523:240-4.

21. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES and Dekker J. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009;326:289-93.

22. Perrin MJ, Adler A, Green S, Al-Zoughool F, Doroshenko P, Orr N, Uppal S, Healey JS, Birnie D, Sanatani S, Gardner M, Champagne J, Simpson C, Ahmad K, van den Berg MP, Chauhan V, Backx PH, van Tintelen JP, Krahn AD and Gollob MH. Evaluation of genes encoding for the transient outward current (Ito) identifies the KCND2 gene as a cause of J-wave syndrome associated with sudden cardiac death. *Circ Cardiovasc Genet*. 2014;7:782-9.

23. Gao C, Ren S, Lee JH, Qiu J, Chapski DJ, Rau CD, Zhou Y, Abdellatif M, Nakano A, Vondriska TM, Xiao X, Fu XD, Chen JN and Wang Y. RBFox1-mediated RNA splicing regulates cardiac hypertrophy and heart failure. *J Clin Invest*. 2016;126:195-206.

24. Martin JF, Schwarz JJ and Olson EN. Myocyte enhancer factor (MEF) 2C: a tissue-restricted member of the MEF-2 family of transcription factors. *Proceedings of the National Academy of Sciences of the United States of America*. 1993;90:5282-6.

25. Okamoto R, Li Y, Noma K, Hiroi Y, Liu PY, Taniguchi M, Ito M and Liao JK. FHL2 prevents cardiac hypertrophy in mice with cardiac-specific deletion of ROCK2. *FASEB J*. 2013;27:1439-49.

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

27. Bickmore WA and van Steensel B. Genome architecture: domain organization of interphase chromosomes. *Cell*. 2013;152:1270-84.

28. Greco CM, Kunderfranco P, Rubino M, Larcher V, Carullo P, Anselmo A, Kurz K, Carell T, Angius A, Latronico MV, Papait R and Condorelli G. DNA hydroxymethylation controls cardiomyocyte gene expression in development and hypertrophy. *Nature communications*. 2016;7:12418.

29. Movassagh M, Choy MK, Knowles DA, Cordeddu L, Haider S, Down T, Siggens L, Vujic A, Simeoni I, Penkett C, Goddard M, Lio P, Bennett MR and Foo RS. Distinct epigenomic features in end-stage failing human hearts. *Circulation*. 2011;124:2411-22.

30. Zhang QJ, Chen HZ, Wang L, Liu DP, Hill JA and Liu ZP. The histone trimethyllysine demethylase JMJD2A promotes cardiac hypertrophy in response to hypertrophic stimuli in mice. *The Journal of clinical investigation*. 2011;121:2447-56.

31. Chang CP and Han P. Epigenetic and lncRNA regulation of cardiac pathophysiology. *Biochim Biophys Acta*. 2016;1863:1767-71.

32. Dorn GW, 2nd and Matkovich SJ. Epitranscriptional regulation of cardiovascular development and disease. *The Journal of physiology*. 2015;593:1799-808.

33. Sayed D, He M, Yang Z, Lin L and Abdellatif M. Transcriptional Regulation Patterns Revealed by High Resolution Chromatin Immunoprecipitation during Cardiac Hypertrophy. *The Journal of biological chemistry*. 2013;288:2546-58.

34. Giorgetti L, Lajoie BR, Carter AC, Attia M, Zhan Y, Xu J, Chen CJ, Kaplan N, Chang HY, Heard E and Dekker J. Structural organization of the inactive X chromosome in the mouse. *Nature*. 2016;535:575-9.

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

36. Zuin J, Dixon JR, van der Reijden MI, Ye Z, Kolovos P, Brouwer RW, van de Corput MP, van de Werken HJ, Knoch TA, van IWF, Grosveld FG, Ren B and Wendt KS. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells.

Proceedings of the National Academy of Sciences of the United States of America. 2014;111:996-1001.

37. Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA and Bruneau BG. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell*. 2017;169:930-944 e22.

38. Fedoriw AM, Stein P, Svoboda P, Schultz RM and Bartolomei MS. Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. *Science*. 2004;303:238-40.

39. Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, Szalaj P, Trzaskoma P, Magalska A, Wlodarczyk J, Ruszczycki B, Michalski P, Piecuch E, Wang P, Wang D, Tian SZ, Penrad-Mobayed M, Sachs LM, Ruan X, Wei CL, Liu ET, Wilczynski GM, Plewczynski D, Li G and Ruan Y. CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. *Cell*. 2015;163:1611-27.

40. Kung G, Konstantinidis K and Kitsis RN. Programmed necrosis, not apoptosis, in the heart. *Circ Res*. 2011;108:1017-36.

41. Monte E, Rosa-Garrido M, Karbassi E, Chen H, Lopez R, Rau CD, Wang J, Nelson SF, Wu Y, Stefani E, Lusis AJ, Wang Y, Kurdistani SK, Franklin S and Vondriska TM. Reciprocal Regulation of the Cardiac Epigenome by Chromatin Structural Proteins Hmgb and Ctcf: IMPLICATIONS FOR TRANSCRIPTIONAL REGULATION. *The Journal of biological chemistry*. 2016;291:15428-46.

42. Oh S, Oh C and Yoo KH. Functional roles of CTCF in breast cancer. *BMB Rep*. 2017;pii: 3893:PMID: 28648147.

43. Hnisz D, Day DS and Young RA. Insulated Neighborhoods: Structural and Functional Units of Mammalian Gene Control. *Cell*. 2016;167:1188-1200.

44. Spurrell CH, Dickel DE and Visel A. The Ties That Bind: Mapping the Dynamic Enhancer-Promoter Interactome. *Cell*. 2016;167:1163-1166.

45. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES and Aiden EL. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014;159:1665-80.

46. Li H and Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25:1754-60.

47. Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, Dekker J and Mirny LA. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods*. 2012;9:999-1003.

48. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen CJ, Vert JP, Heard E, Dekker J and Barillot E. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol*. 2015;16:259.

49. Bolstad BM, Irizarry RA, Astrand M and Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*. 2003;19:185-93.

50. Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES and Aiden EL. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst*. 2016;3:95-8.

51. Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES and Aiden EL. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst*. 2016;3:99-101.

52. Lun AT, Perry M and Ing-Simmons E. Infrastructure for genomic interactions: Bioconductor classes for Hi-C, ChIA-PET and related experiments. *F1000Res*. 2016;5:950.

53. Ay F, Bailey TL and Noble WS. Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. *Genome Res*. 2014;24:999-1011.

54. Yin T, Cook D and Lawrence M. ggbio: an R package for extending the grammar of graphics for genomic data. *Genome Biol*. 2012;13:R77.

55. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R and Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14:R36.

56. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ and Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28:511-5.

57. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL and Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol*. 2013;31:46-53.

58. Goff L, Trapnell C and Kelley D. cummeRbund: Ananlysis, exploration, manipulation and visualization of Cufflinks high-throughput sequencing data. *R package*

version *2160*. 2013;https://bioconductor.org/packages/release/bioc/html/cummeRbund.html.

59. Langmead B and Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357-9.

60. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R and Genome Project Data Processing S. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25:2078-9.

61. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W and Liu XS. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9:R137.

62. Yu G, Wang LG and He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*. 2015;31:2382-3.

63. Shen L, Shao N, Liu X and Nestler E. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics*. 2014;15:284.

64. Carlson M. KEGG.db: A set of annotation maps for KEGG. *R package version 323*. 2016;http://bioconductor.org/packages/release/data/annotation/html/KEGG.db.html.

65. O'Connell TD, Rodrigo MC and Simpson PC. Isolation and culture of adult mouse cardiac myocytes. *Methods Mol Biol*. 2007;357:271-296.

66. Heath H, Ribeiro de Almeida C, Sleutels F, Dingjan G, van de Nobelen S, Jonkers I, Ling KW, Gribnau J, Renkawitz R, Grosveld F, Hendriks RW and Galjart N. CTCF regulates cell cycle progression of alphabeta T cells in the thymus. *EMBO J*. 2008;27:2839-50.

67. Sohal DS, Nghiem M, Crackower MA, Witt SA, Kimball TR, Tymitz KM, Penninger JM and Molkentin JD. Temporally regulated and tissue-specific gene manipulations in the adult and embryonic heart using a tamoxifen-inducible Cre protein. *Circ Res*. 2001;89:20-5.

68. Rau CD, Wang J, Avetisyan R, Romay M, Ren S, Wang Y and Lusis AJ. Mapping genetic contributions to cardiac pathology induced by beta-adrenergic stimulation in mice. *Circulation Cardiovascular Genetics*. 2015;8:40-49.

Chapter 3: Spatial Principles of Chromatin Architecture Associated with Organ-Specific Gene Regulation

Douglas J. Chapski, Manuel Rosa-Garrido, Nan Hua,

Frank Alber, Thomas M. Vondriska

[This research was originally published by Chapski *et al.* Spatial Principles of Chromatin Architecture Associated with Organ-Specific Gene Regulation. *Frontiers in Cardiovascular Medicine*. 2019 Jan 15;5:186. PMID: 30697540 © Chapski, Rosa-Garrido, Hua, Alber and Vondriska.]

Keywords: transcription, chromatin conformation capture, genomics, chromatin structure, epigenetics

ABSTRACT

Packaging of the genome in the nucleus is a nonrandom process that is thought to directly contribute to cell type-specific transcriptomes, although this hypothesis remains untested. Epigenome architecture, as assayed by chromatin conformation capture techniques such as Hi-C, has recently been described in the mammalian cardiac myocyte and found to be remodeled in the setting of heart failure. In the present study, we sought to determine whether the structural features of the epigenome are conserved between different cell types by investigating Hi-C and RNA-seq data from heart and liver. Investigation of genes with enriched expression in heart or liver revealed nuanced interaction paradigms between organs: first, the $\log_2$ ratios of heart:liver (or liver:heart) intrachromosomal interactions are higher in organ-specific gene sets (p = 0.009), suggesting that organ-specific genes have specialized chromatin structural features. Despite similar number of total interactions between cell types, intrachromosomal interaction profiles in heart but not liver demonstrate that genes forming promoter-to-transcription-end-site loops in the cardiac nucleus tend to be involved in cardiac-related pathways. The same analysis revealed an analogous organ-specific interaction profile for liver-specific loop genes. Investigation of A/B compartmentalization (marker of chromatin accessibility) revealed that in the heart, 66.7% of cardiac-specific genes are in compartment A, while 66.1% of liver-specific genes are found in compartment B, suggesting that there exists a cardiac chromatin topology that allows for expression of cardiac genes. Analyses of interchromosomal interactions revealed a relationship between interchromosomal interaction count and organ-specific gene localization ($p = 2.2 \times 10^{-16}$) and that, for both organs, regions of active or inactive chromatin tend to segregate

in 3D space (i.e. active with active, inactive with inactive). 3D models of topologically associating domains (TADs) suggest that TADs tend to interact with regions of similar compartmentalization across chromosomes, revealing *trans* structural interactions contributing to genomic compartmentalization at distinct structural scales. These models reveal discordant nuclear compaction strategies, with heart packaging compartment A genes preferentially towards the center of the nucleus and liver exhibiting preferential arrangement towards the periphery. Taken together, our data suggest that intra- and interchromosomal chromatin architecture plays a role in orchestrating tissue-specific gene expression.

INTRODUCTION

Before DNA was universally recognized as the genetic material, it was hypothesized that nuclear proteins may be responsible for how the same DNA does different things in the various cell types of a multicellular organism.[1] Since around the same time, it has been appreciated that nuclear proteins, histones in particular, exhibit distinct biochemical properties across cell types and stages of development[2]—DNA itself has long been known to be modified by methylation according to similar physiological variables.[3] In the ensuing decades, it has become clear that histone modification and nucleosome positioning play a central role in specifying distinct transcriptomes,[4,5] but the implications for chromatin structure have remained uncertain.

More recently, the emergence of chromatin capture technology combined with next generation sequencing has enabled unprecedented analyses of endogenous chromatin structure with increasing levels of resolution.[6-8] Chromatin compartmentalization has been characterized as an intrinsic property of nuclear architecture, denoting regions tending to be more accessible as “compartment A” and those less accessible “compartment B.”[7] In addition to compartmentalization, Hi-C data can reveal properties of chromatin looping.[9,10] Putative gene loops have also been identified from RNA Polymerase II ChIP-seq datasets,[11] wherein genes have their promoters and transcription end sites in close 3D proximity to facilitate continued transcription. Folding of the genome is a non-random, reproducible process that favors local over long range interactions. This behavior leads to the formation of topologically associating domains (TADs), which exhibit greater interactions within themselves than between, constituting a structural unit greater in scale than the nucleosome (TADs are composed of kilobases of DNA and associated

nucleosomes) and smaller than the chromosome, with boundary regions between TADs being ostensibly responsible for cordoning distinct regions of transcriptional behavior. HiC, one of the principle techniques for genome wide chromatin structural analysis, has now been deployed in multiple laboratories around the world, as well as in multiple cell types, revealing TADs and chromatin compartmentalization to be conserved structural rules governing genome organization.[12,13]

These observations raise the following question: if TADs are a conserved feature of epigenomes across cell types, where does the specificity in structure arise? Compounding this question is the fact that, until recently, chromatin conformation capture studies have been often carried out in either cell culture or whole tissue extracts, making it possible to evaluate neither terminally differentiated cells nor the cell type-specific nature of chromatin structure. While we understand transcriptome changes across multiple organs and disease states, a major gap in our basic understanding of organ function is how genome architecture varies between cells and how this relates to gene expression.

To address these gaps in knowledge, we investigated chromatin structural differences between heart and liver, and how they relate to tissue-specific gene expression programs. Specifically, we studied the role of genomic interactions (both intra- and interchromosomal) in organ-specific gene architecture. The analysis reveals a concordance between interaction frequency and organ-specific gene expression between tissues. We also explored compartment differences between organs, demonstrating that gene expression paradigms in distinct tissues act concertedly with their organ-specific compartmentalization pattern to regulate function of the cell. Lastly, we show that more

interchromosomal interactions exist at organ-specific genes, and that about half of such interactions bridge distinct compartments within both cardiac and liver nuclei. Together, these investigations reveal organ-specific chromatin conformations that may contribute to cell identity in heart and liver.

MATERIALS AND METHODS

Hi-C Bioinformatics

Hi-C datasets from this study were downloaded from NCBI GEO: Isolated cardiac myocyte data[14] were downloaded using accession number GSE96693 (Control_HiC). Liver data (acquired from isolated hepatocyte nuclei)[15] were downloaded using accession number GSE104129 (Hi-C reps1-5). This dataset comes from wild-type C57BL/6J mice whose hepatocyte nuclei were isolated via homogenization and then crosslinked in 1% formaldehyde in PBS and quenched in glycine (125 mM final concentration) for Hi-C. Hi-C libraries for both datasets were generated using Hi-C protocols based on[9] with small changes described in previous work.[14,15] Libraries for both heart and liver were constructed using MboI as the restriction endonuclease and sequenced deeply enough to achieve 5kb resolution contact matrices (see Table 3-1 for sequencing depth and valid interaction pair numbers, as determined by our pipeline described below).

Hi-C datasets were run through the HiC-Pro analysis pipeline,[16] version 2.10.0. Briefly, raw FASTQ files from all biological replicates were combined for each organ, and aligned to mm10 using an MboI-digested restriction fragment list generated by HiC-Pro. After the quality_checks step, we built 5kb resolution contact maps and performed iterative correction and eigenvector decomposition (ICE normalization, first described in[17]), using HiC-Pro. We also built 100kb contact maps for 3D model building. We then converted contact matrices to a Fit-Hi-C[10] friendly format using the hicpro2fithic.py Python script provided with HiC-Pro, with the raw contact matrices and ICE biases as inputs. Fit-Hi-C version 2.0.3 was used to determine significant intrachromosomal interactions, using

the following parameters: -b 200 -r 5000 -p 2. The advantage of Fit-Hi-C version 2 (as compared to version 1) is that it can determine significant intrachromosomal interactions without constraining the data to mid-range distances. That is, we can use Fit-Hi-C version 2 to identify regions of significant intrachromosomal interaction along *entire* chromosomes, at 5kb resolution. Another benefit of Fit-Hi-C is that it reports a q-value for each interaction, and we can filter for significant ($q < 0.01$) ones. For significant interchromosomal interaction identification, we performed a similar Fit-Hi-C analysis, but with the following parameters to investigate interactions that are not on the same chromosome: -b 200 -r 5000 -x interOnly. Significant intra- and interchromosomal interactions ($q < 0.01$) at 5kb resolution are quantified in Table 3-1. A/B compartmentalization was calculated on 5kb resolution contact matrices.[14] For each bin in the genome, differences in A/B compartmentalization between heart and liver were noted and shown in Figure 3-3C. All analyses in this study were done on autosomes only, unless otherwise stated.

To generate 3D models of topologically associating domains, we first ran TopDom[18] version 0.0.2 on ICE-normalized 100kb matrices to generate a list of TADs in cardiac and liver Hi-C data, using *window.size* = 3 as a parameter. We then used Population-based Genome Structure (PGS) software[19] to generate 10,000 3D models of the genome (autosomes + chrX), using default parameters, the 100kb matrices, and TAD calls as inputs. The contact probabilities between TADs in the resulting population of genome structures are statistically consistent with the contact probability matrix from Hi-C experiments (Figure 3-10). This resulted in an extracted list of *xyz* coordinates for each TAD that were used to generate PDB files for visualization, structure analysis, as well as

distance matrix calculations to determine the closest interchromosomal TADs for each TAD in the genome (custom R scripts). For both heart and liver data, 100kb resolution TADs were designated as being in compartment A or B based on the majority compartmentalization status of the 5kb bins (A/B analysis described in paragraph above) that lie within each 100kb resolution TAD.

RNA-seq Bioinformatics

RNA-seq data from this study were downloaded from NCBI GEO: Isolated cardiac myocyte data corresponding to our previous study[14] were downloaded using accession number GSE96693 (Control_RNAseq Replicates 1-3). Mouse (C57BL/6) whole liver RNA-seq data from ENCODE Portal were downloaded using ENCODE Data Coordination Center accession number ENCSR000BYS (which is identical to the data at NCBI GEO accession number GSE90180). Raw FASTQ files from two biological replicates of liver tissue were downloaded. For both cardiac and liver RNA-seq library prep, rRNA was depleted and polyA selection performed. For the bioinformatics analysis, raw paired-end FASTQ files were aligned to the mm10 reference genome (Ensembl release 81) using HISAT2[20] version 2.1.0 with an mm10 HISAT2 index (built in-house). Resulting SAM alignments were converted to BAM format and sorted by name with Samtools[21] version 1.7. Gene counts were determined using htseq-count[22] version 0.9.1, with the sorted BAM alignments and a GTF of known Ensembl genes from release 81 as input. The Bioconductor package DESeq2[23] was then used to pre-filter the genes that have at least 10 reads between any of the replicates (3 heart and 2 liver) and to collapse replicates by organ with the collapseReplicates() function. Then, the counts() function with the

normalized=T option resulted in a single normalized count value for all genes for each organ, for use during downstream analyses.

Promoter-TES Analysis

Genes with promoter-TES loops were identified by determining, using the Bioconductor package InteractionSet,[24] the genes that have significant ($q < 0.01$) Fit-Hi-C interactions with the promoter (-2000 to +200bp form TSS) *and* the TES of a gene. Genes that have such interactions underwent KEGG analysis using KEGG.db,[25] a package in Bioconductor, with custom graphics generation using ggplot2 in R. Indicated p-values are calculated using a hypergeometric test.

Organ-Specific Gene Designation

We designated organ-specific genes using the Human Protein Atlas,[26] specifically the subset of genes that are enriched in heart and liver at the mRNA level. The Human Protein Atlas defines “tissue enriched” genes as having at least 5-fold higher mRNA expression in the organ of interest when contrasting against all other organs.[26] Human Ensembl gene identifiers from these tables were fed into biomaRt[27] in R and converted to Mouse Ensembl identifiers. For further analysis, we filtered to keep gene coordinates on murine autosomes.

RESULTS

Chromatin Microenvironments Facilitate Organ-Specific Gene Interaction

To examine whether nuclei of different cells create chromatin micro-environments for the transcription profiles they produce, we first designated cardiac- and liver-specific genes as those having 5X higher expression in the organ of interest when compared to all other tissues in the Human Protein Atlas.[26] We then examined the chromatin interactions, detected in cardiomyocyte or hepatocyte Hi-C experiments, around these organ specific genes.

Both datasets were sequenced to a similar depth (~1.3-1.5 billion read pairs; Table 3-1) and achieved a similar number of significant (q < 0.01) intrachromosomal Fit-Hi-C interactions (115,843 in heart and 90,587 in liver; Table 3-1). We quantified the $\log_2$ ratio of cardiac/liver Fit-Hi-C intrachromosomal interactions at cardiac or liver gene loci and found that in both organs, there was a greater ratio of interactions around that organ's specific genes (Figure 3-1A; p = 0.009). These findings suggest that structural organization in 3D underpins cell type specific transcriptomes through greater frequency of interactions (Figure 3-1B).

Organ-Specific Compartmentalization Governs Heart and Liver mRNA Expression

To understand the accessibility of cardiac- and liver-specific genes within the context of heart and liver chromatin, we calculated the A/B compartmentalization status of these genes as determined from Hi-C experiments. In the cardiac Hi-C data, the majority (66.7%) of cardiac-specific genes are found in compartment A (the accessible

compartment), while the majority of liver-specific genes (61.9%) are found in compartment B (the less accessible compartment) (Figure 3-2A, *left*). Contrastingly, the majority of *both* cardiac- and liver-specific genes are found in compartment A in the liver Hi-C data (Figure 3-2A, *right*; 63.0% of cardiac and 63.6% of liver genes). As a positive control for the gene selections strategy, cardiac- and liver-specific genes are more highly expressed at the mRNA level in heart and liver cells, respectively (Figure 3-2B) in the experiments used for this study. Cardiac-specific genes in compartment A are more highly expressed than those in compartment B ($p = 1.4 \times 10^{-10}$ between heart and liver for genes in compartment A, $p = 9.6 \times 10^{-25}$ for genes in compartment B), and the same is true for liver-specific genes in liver tissue ($p = 9.3 \times 10^{-27}$ between liver and heart genes in compartment A, $p = 2.6 \times 10^{-7}$ for genes in compartment B; Figure 3-2B). Taken together, these data suggest that the heart contains cardiac-specific chromatin conformations that allow for cardiac (and not liver) gene accessibility and expression via a more open compartmentalization regime at specific cardiac gene loci. In contrast, liver chromatin can tolerate more cardiac specific genes in active compartments, whereas the reverse is not true for liver genes in cardiac chromatin.

Interaction Profiles and Compartmentalization of Genes in 3D

Cardiac- and liver-specific genes have increased accessibility and a larger number of intrachromosomal interactions at organ-specific genes. However, the distribution of intrachromosomal interactions across genomic features, as well as the compartment change of 5kb bins in the genome, could contribute to this phenomenon in both organs. To investigate whether intrachromosomal interactions in heart and liver have different

localization across genomic features (promoters, exons, introns, intergenic regions), we performed an overlap of significant intrachromosomal Fit-Hi-C interaction anchors (i.e. one side of an interaction pair) with these regions. Notably, in both heart and liver, we observe an almost identical distribution of Fit-Hi-C anchors (Figures 3-3A, 3-3B). Intrachromosomal Fit-Hi-C interactions are enriched within promoters, exons, and introns, and depleted from introns and intergenic regions (Figure 3-3B), suggesting a common packaging logic characterized by increased interactions in regions that contain genes.

To understand how A/B compartmentalization differs between heart and liver, we determined which 5kb bins of the genome have a difference in compartment status between heart and liver. Five percent of bins are in compartment B in the heart and A in the liver, while 7% of bins are in compartment A in the heart and B in the liver (Figure 3-3C), for a total of 12% of the genome that shows compartmentalization differences between both organs. Genes that are in compartment A in the heart but are in compartment B in the liver are more highly expressed at the mRNA level in the heart than in the liver ($p = 5.2 \times 10^{-9}$, Figure 3-3D). Contrastingly, genes that are in compartment B in the heart and A in the liver are more highly expressed in the liver than in the cardiac RNA-seq data ($p = 2.7 \times 10^{-23}$, Figure 3-3D). Taken together, these data suggest that chromatin organization directly contributes to organ-specific gene regulation at a global scale.

Genes with Promoter-TES Interactions Are Organ-Specific

We next sought to determine whether there are organ-specific gene loops that govern cardiac- or liver-specific organ function. Here we define gene loops as significant

(q < 0.01) intrachromosomal Fit-Hi-C interactions that overlap both the promoter region (-2000 to +200 bp from transcription start site) and the transcription end site (TES) of a gene (Figure 3-4A). Our analyses revealed 492 and 298 genes (overlap = 78) with promoter-TES looping in the cardiac and liver Fit-Hi-C data, respectively (Tables 3-2 and 3-3; note this analysis was unbiased—genes were not preselected for organ specific functions as in preceding analyses). KEGG pathway analysis on cardiac loop genes reveals enrichments for terms such as dilated cardiomyopathy, vasopressin-regulated water reabsorption, and hypertrophic cardiomyopathy (Figure 3-4B) (the genes within these terms include: *Adcy6*, *Aqp3*, *Creb3l4*, *Des*, *Itgb5*, *Itga9*, *Myl2*, *Myl3*, and *Stx4a*). The same analysis on liver gene loops reveals enrichments for phenylalanine, tyrosine and tryptophan biosynthesis, phenylalanine metabolism, allograft rejection, and tryptophan metabolism (Figure 3-4C) (the genes of which include: *Cyp1a1*, *Cyp1a2*, *Fasl*, *Got1*, *H2-T10*, *Il2*, *Il12a*, *Lao1*, and *Tat*).

Interchromosomal Interactions Have Different Compartment Status in Heart and Liver

Examination of interchromosomal interactions allows for exploration of regional apposition—and potentially regulation—between distinct chromosomes. Significant (q < 0.01) interchromosomal interactions were identified in the cardiac and liver Hi-C datasets and are summarized in Table 3-1. To determine whether interchromosomal interactions from the cardiac Hi-C data preferentially overlap cardiac specific genes, we overlapped these regions with the cardiac- and liver-specific genes from the analyses in Figures 3-1 to 3-3. In cardiac chromatin, 540 interchromosomal interactions overlap with cardiac-

specific genes, while only 63 interactions overlap with liver-specific genes. In the liver chromatin, 433 interchromosomal interactions overlap with liver-specific genes, whereas only 243 overlap with cardiac-specific genes. These data suggest that interchromosomal interactions at organ-specific genes depend on the nuclear environment within the organ of interest (p = 2.2 x 10^{-16}, Fisher's exact test; Figure 3-5). To confirm this observation, we performed a simulation which resulted in no relationship between randomly selected genes from the genome and interchromosomal Fit-Hi-C interactions in either organ (p = 1, Fisher's exact test; Figure 3-5). To investigate the compartmentalization of interchromosomal interactions in cardiac and liver nuclei, we determined the compartment status at each anchor of these interactions (Figure 3-6A). In the cardiac Hi-C data, 7,884 significant interchromosomal interactions have both ends in compartment A, while 20,151 have both ends in compartment B, and 23,335 have each end in a different compartment (Figure 3-6B). In the liver Hi-C data, 14,466 significant interchromosomal interactions have both ends in compartment A, while 40,071 have both ends in compartment B, and 39,764 have each end in a different compartment (Figure 3-6B). In heart and liver Hi-C data, 45% and 42% of significant interchromosomal Fit-Hi-C interactions, respectively, have one end in compartment A and the other in compartment B. Stated another way, about half of significant interchromosomal interactions extend to other compartments, while the other half share compartment status. This observation suggests the existence of chromatin regions that localize to the same area in the nucleus and yet exhibit distinct compartmentalization and potentially distinct accessibility features.

HiC data is informative to define regions of local interaction, but how these substructures of the epigenome arrange in 3D has remained an open question. We

performed 3D reconstruction of cardiac and liver epigenomes based on HiC data, establishing models for how chromosomes fold and for how they associate with other chromosomes, using PGS.[19] The approach generates a large population of 3D genome structures, in which TAD domains are represented by spheres and are then packed into the nucleus in such a way that the formation of contacts between TAD domains is statistically consistent with the contact probability matrix from Hi-C experiments (Figure 3-10). These models reveal distinct chromosomal structures within liver or cardiac epigenomes (i.e. allowing comparison of one chromosome to another), enable comparison of the individual chromosomes between organs, elucidate the surfaces of interaction between chromosomes (Figure 3-7, Online Movie 1) and reveal insights into the spatial organization of chromatin compartments.

To investigate the distribution of different chromatin features within the nuclear space, we divided the nuclear volume into 5 concentric shells in such a way that each shell contains 20% of the total number of TADs per structure. Based on their radial positions, all TADs in each of the 10,000 genome structures are then partitioned into the 5 shells. We then measured the probability for a TAD in a given subcompartment (A/B) to be localized in each of the concentric shells (Figure 3-8A). We observe striking differences in the internal organization of the compartments. In heart cells, chromatin in compartment A shows the highest localization probability in the most inner shells (shell 1 in Figure 3-8A), and the probability gradually decreases towards the outer regions (shell 5 in Figure 3-8A, top left panel). This observation is consistent with previous observations that showed highly transcribed genes to be localized towards the interior regions of the nucleus.[28] Compartment B shows the opposite behavior, with the highest localization

probability for the outer most shell (Figure 3-8A, lower left panel), consistent with the location of heterochromatin and lamina associated domains at the nuclear envelope.[29,30] In contrast, liver cells show a different spatial organization in the models. Compartment A is more evenly distributed with the highest localization probability at the outermost shell, while compartment B shows a slight decrease in localization probability towards the most outer shell.

To gain a quantitative understanding of interchromosomal TAD-TAD colocalization, we studied the compartment composition at the interchromosomal boundaries. At each TAD position, we determined all TADs that are localized within a distance of 500 nm and are part of a different chromosome. We then determined the percentage of A/B compartment found in this group of inter-chromosomal TAD neighbors. The heart genome shows a high preference for TADs in the same chromatin compartment across chromosome boundaries, indicating a high level of compartmentalization across chromosome borders. In liver cells, we observe a different behavior. While TADs in subcompartment B also show a high preference to be in proximity to TADs in same state, TADs of state A do not show a preference for the same state, showcasing the different global organization of the genome in liver nuclei.

We also calculated the average radial position of each TAD with respect to the nuclear center (Figure 3-8B; see Figure 3-9 for comparison of all chromosomes). When plotting the average radial positions for each TAD across a chromosome we observe distinct regional differences with well-defined local minima and maxima (Figure 3-8B). TADs corresponding to minima are on average more interior located than directly neighboring regions in the same chromosome. These radial position profiles are markedly

different for the same chromosomes in the two tissues. The correlation between the radial position profiles is very low, and in some regions even anti-correlated (Figure 3-8B). These distinctions are further illustrated when examining the likelihood of regions from the same compartment to interact with each other (Figure 3-8C).

Finally, we examined the localization of chromatin from a gene centric view, determining the relative positioning of heart and liver specific genes in the different nuclei. In agreement with the observations from Figure 3-8A, this gene centric analysis revealed a preference of interior localization of genes in cardiac nuclei and the antithetical behavior in liver nuclei. In summary, our structure-based calculations support the notion that, on a TAD scale (hundreds of kilobases), there are major structural differences in the global structural organization of liver and heart genomes.

DISCUSSION

How chromatin structure underpins gene expression has ramifications across biology and medicine. In the cardiovascular realm, as in other areas of epigenomics research, this question has largely been answered from the perspective of histone modifications,[31,32] enhancers,[33] chromatin remodeling enzymes,[34] transcription factors,[35] DNA methylation[36] and, more recently, long noncoding RNAs.[37] Lacking from all of these studies has been a direct measurement of chromatin structure, rather than relying on implications of structure and accessibility as a result of the actions of other proteins or modifications. Recent chromatin conformation capture experiments[14,38] in human and mouse cardiomyocytes now make possible examination of cardiac chromatin structure and investigation of how this structure contributes to lineage specification and heart disease.

The current study demonstrates that organ-specific genes preferentially localize in 3D in the nuclei of the organs in which they are transcribed. This conceptually straightforward hypothesis has never, to our knowledge, been tested experimentally and reveals a structural underpinning for cell type-specific transcriptomes. These observations also support the concept of transcriptional neighborhoods,[39] or transcription factories, which have been hypothesized to coordinate RNA production from a select subset of DNA templates but which has never been tested in cardiovascular cell types. A caveat arising from the data used for this study (cardiac HiC and RNA-seq data were from isolated adult mouse cardiac myocytes; liver HiC data were from isolated hepatocytes and RNA-seq data from whole tissue) is that some of the cell type-specific differences in hepatocyte gene expression may be obfuscated by other cells present in the entire liver,

although this should have no bearing on the analyses of chromatin architecture, which in each case were performed on an isolated cell population from adult C57BL/6J mice. Because the primary data used for these analyses were collected in two different laboratories, there is a concern that the differences in genomic organization may be attributable to confounding variables unrelated to the cell type differences. Mitigating this concern is the fact that the animals were the same genetic strain, housed in similar environments and sacrificed at the same time of day. Moreover, the sequencing data enabled identification of a comparable number of total interaction pairs in cardiac (807,707,536) and liver (701,407,381) experiments, producing interactions maps at comparable resolution (~5kb).

Our comparison of liver and cardiac chromatin structure reveals widespread differences in compartmentalization, some but not all of which coordinate with transcriptional behaviors that vary between the organs. This finding is intriguing, given the fact that altered compartmentalization following the development of pressure overload-induced cardiac hypertrophy and failure is very minor:[14] localization of genes within organ specific chromatin scaffolds is specific to cell type and resilient against pathophysiological stress. It is tempting to speculate that the differences in chromatin architecture may reflect the more proliferative nature of the liver compared to the heart. Hepatocytes, like cardiomyocytes, are terminally differentiated, and the majority of these cells—in a healthy, unstressed liver—would not be actively undergoing mitosis (and the associated genomic rearrangements). However, the liver has a well established ability to regenerate upon physical damage and/or cell death. Perhaps the liver prepares for such an eventuality by allowing a greater number of genes to exist in accessible regions of

chromatin, although further experiments will be necessary to provide evidence for this conjecture, including examination of chromatin architecture in proliferative liver tissue.

The results of the analysis of gene looping data were particularly revealing: heart and liver establish comparable numbers of promoter to TES gene loops, however this specific class of loops appears in different genes in the different organs. These findings support that at multiple scales, including the level of gene looping in addition to compartmentalization as mentioned above, structural organization of the epigenome is cell type specific.

The majority of chromatin conformation capture studies that have emerged the past few years have focused exclusively on intrachromosomal interactions. The adult cardiac myocyte, which does not divide, is an interesting test case to explore the role of interchromosomal contact surfaces in genome function—principally, although not exclusively, via gene regulation. A liver HiC dataset of comparable sequencing depth afforded the opportunity to explore contrasting features of such interactions, should they exist, within the same genome housed in separate cells' nuclei. Both epigenomes exhibited similar levels or interchromosomal interactions and in both cases, they were enriched in genes associated with the function of that cell type. Combining these interactions with 3D renderings of genomes in heart and liver provided a unique opportunity to investigate differences in chromosome folding and nuclear organization. Several observations emerged: liver and heart cells not only package their genomes differently, but they appear to obey distinct general principles of organization, wherein heart genomes preferentially localize compartment A regions towards the center and compartment B regions towards the periphery, whereas liver cells do not exhibit this

behavior. Future studies will investigate whether interchromosomal interaction surfaces participate in such behaviors as cell proliferation, whether they change with age or are dependent on developmental state, and what non-DNA molecules inhabit the surfaces of interchromosomal apposition, presumably orchestrating the reproducible formation of these structures.

ACKNOWLEDGEMENTS

We thank Dr. Christoph D. Rau and members of the Vondriska Lab for helpful discussions. We also thank Dr. Ferhat Ay for assistance with use of the Fit-Hi-C algorithm.

AUTHOR CONTRIBUTIONS

DJC and TMV conceived the study. DJC performed bioinformatics and statistical analyses. DJC and MRG generated figures and diagrams. NH and FA generated and analyzed 3D genomic models. DJC and TMV wrote the paper. All authors approved the content of the manuscript.

FUNDING

The Vondriska Lab is supported by grants from the National Institutes of Health (R01 HL129639, R01 HL105699, R01 HL143058 to TMV) and the Cardiovascular Theme in the David Geffen School of Medicine at UCLA.

CONFLICT OF INTEREST STATEMENT

The authors have nothing to disclose.

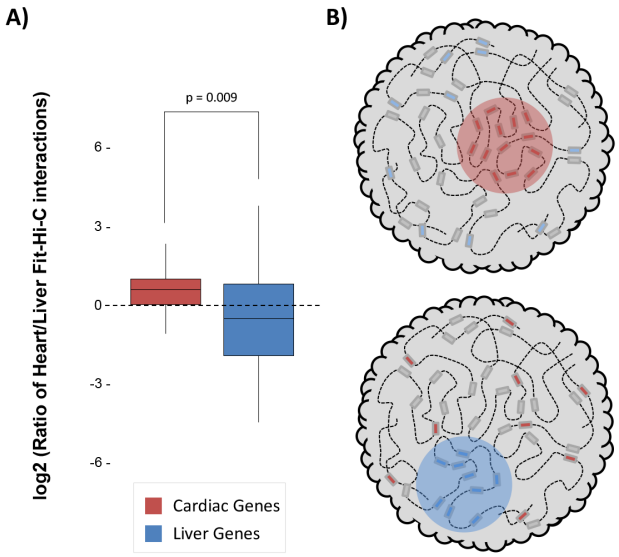


Figure 3-1. Significant Fit-Hi-C interactions are observed in organ-specific genes. (A) The log2 ratio of significant ($q < 0.01$) Fit-Hi-C interactions (Heart/Liver) in cardiac-specific genes (red) is higher than in liver-specific genes (blue). (B) Schematic demonstrating the hypothesis that regions of the nucleus containing heart-specific genes (left, red circle) contain more significant Fit-Hi-C interactions in cardiac Hi-C data when compared to liver Hi-C data (right panel and blue circle show same principle for liver genes in the liver nucleus).

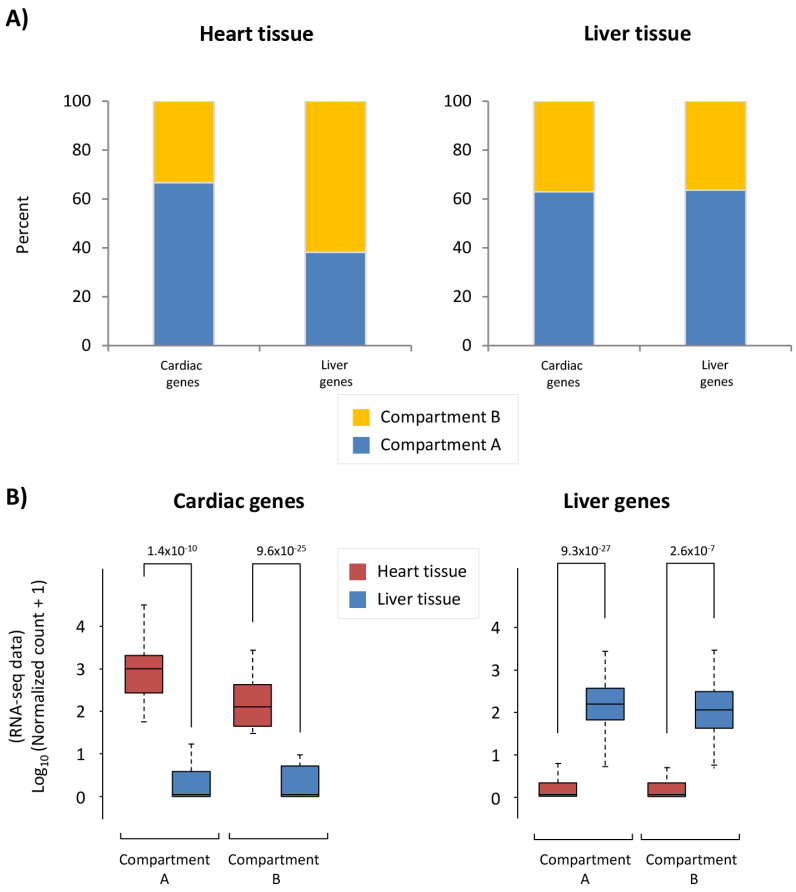


Figure 3-2. Cardiac and liver genes are situated in open compartments. (A) Left, in the cardiac Hi-C data, the majority of cardiac-specific genes are found in compartment A (more accessible), whereas the majority of liver genes are in compartment B (less accessible). Right, the majority of liver genes are found in compartment A in the liver Hi-C data; this is also the case for cardiac genes in liver Hi-C data. (B) Left, Cardiac-specific genes are more highly expressed in heart than liver genes. Right, liver-specific genes are

more highly expressed in liver tissue than in the heart. The y-axis shows $\log_{10}$ of the normalized RNA-seq read counts, which are calculated according to the DESeq2 read count normalization method for each gene (see methods). Indicated p-values were calculated using Wilcoxon rank-sum test. Note an observed trend of higher expression for those genes that lie within compartment A when compared to those in compartment B.

A)

	Number of intrachromosomal interactions						% Percent			
	Interactions	Anchors	Promoter	Exon	Intron	Intergenic	Promoter	Exon	Intron	Intergenic
Heart	115843	231686	42558	45197	45157	98774	18.37	19.51	19.49	42.63
Liver	90587	181264	30377	32480	35793	82614	16.76	17.92	19.75	45.58

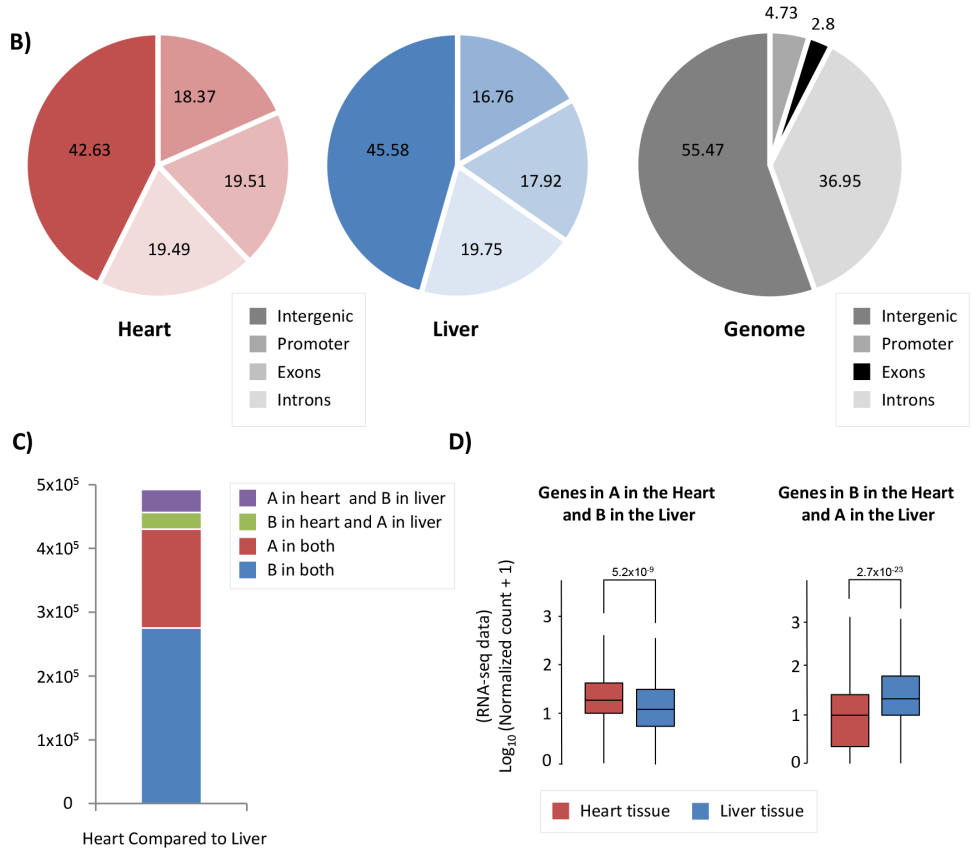


Figure 3-3. Significant Fit-Hi-C interactions have a similar distribution among annotated features of the genome. (A) Table describing promoter/exon/intron/intergenic distribution of Fit-Hi-C interaction anchors. The number of anchors is double the number of significant ($q < 0.01$) Fit-Hi-C interactions identified in the study. (B) Pie charts depicting the data shown in (A). (C) Compartmentalization

differences between heart and liver reveal that ~5% of the genome is in a different compartment between heart and liver. (D) The genes in compartment A in the heart and B in the liver tend to have higher expression in the heart (left, red box) than in the liver (left, blue box); in contrast, genes in compartment B in the heart and A in the liver are more highly expressed in the liver. The y-axis shows $\log_{10}$ of the normalized RNA-seq read counts, which are calculated according to the DESeq2 read count normalization method for each gene. Indicated p-values are calculated using Wilcoxon rank-sum test.

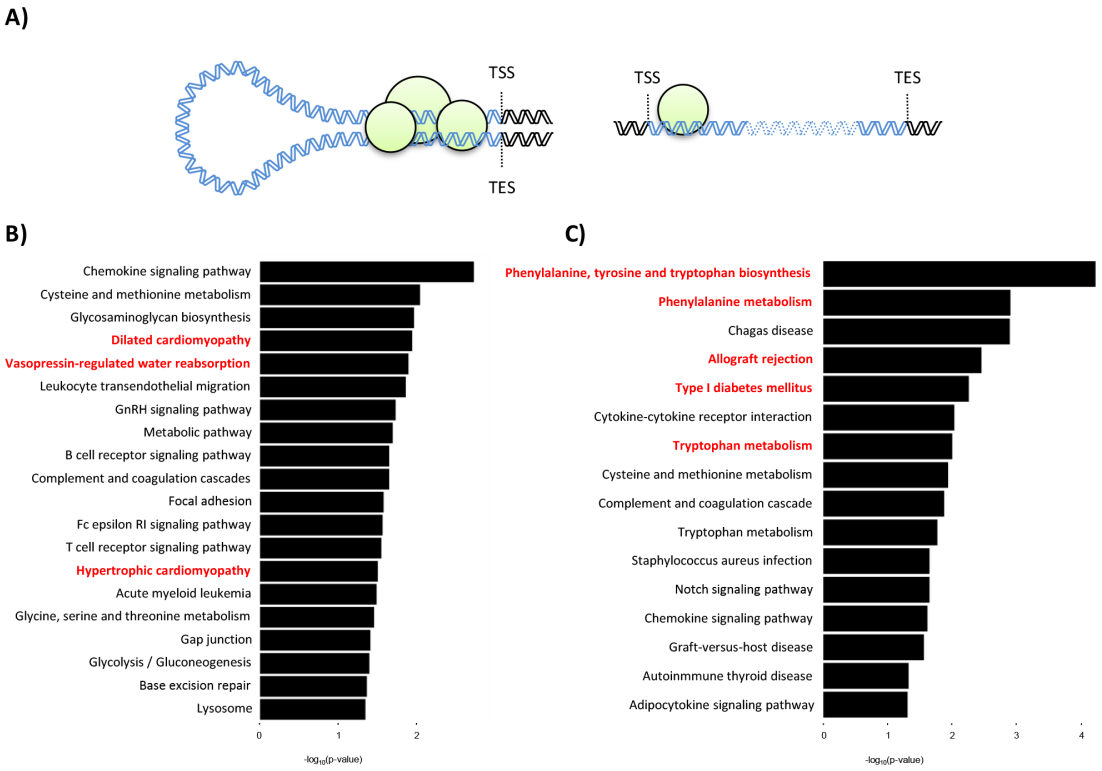


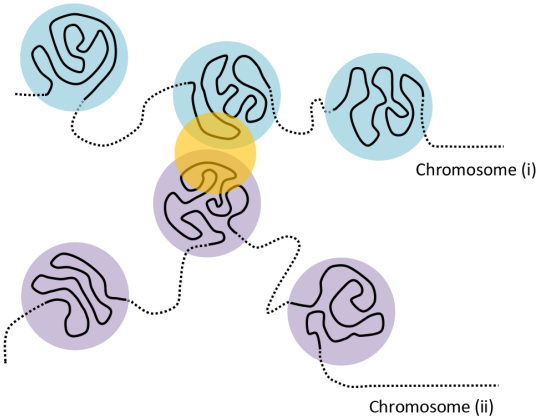
Figure 3-4. Genes with promoter-TES loops tend to have organ-specific functions within the cell. (A) Illustration depicting a gene with promoter-TES interactions (left) and a gene with no promoter-TES interactions (right). (B) KEGG pathway analysis of the genes with significant ($q < 0.01$) Fit-Hi-C interactions between promoters and transcription end sites in the cardiac Hi-C data. Cardiac-related terms are highlighted in red. (C) KEGG

pathway analysis for genes with significant promoter-TES Fit-Hi-C interactions in liver Hi-C data. Liver-related terms are highlighted in red. For panels (B) and (C), p-values are calculated using a hypergeometric test.

	Measured Data		Simulated Data	
	Cardiac genes	Liver genes	Cardiac genes	Liver genes
Cardiac Hi-C data	540	66	39	372
Liver Hi-C data	243	433	72	681

Figure 3-5. Interchromosomal Fit-Hi-C interactions are preferentially found at organ-specific genes. Analysis of interchromosomal Fit-Hi-C interactions ($q < 0.01$) reveals that more interactions are found at cardiac-specific genes in the cardiac Hi-C data, while more interactions are found at liver genes in the liver data (Measured Data, green, $p = 2.2 \times 10\text{-}16$, Fisher's exact test). To calculate the frequency of interactions at random genes, simulations were performed on the cardiac and liver Hi-C data at random genes. Simulations were repeated 10,000 times for each cell in the blue table, and the median number of interactions at random genes was kept for statistical testing (Simulated Data, blue, $p = 1$, Fisher's exact test).

A)



B)

	Cardiac Genome	Liver Genome
Compartment A only	7884	14466
Compartment B only	20151	40071
Between compartments	23335	39764

Figure 3-6. Interchromosomal Fit-Hi-C interactions bridge regions of differing A/B compartmentalization. (A) Concept figure showing interaction (yellow circle) between regions from different chromosomes. The regions on one chromosome (i) are shown in light blue, while regions on the other chromosome (ii) are shown in violet. (B) Analysis of all significant Fit-Hi-C interactions ($q < 0.01$) reveals that approximately half of interchromosomal interactions in both heart and liver nuclei act as a bridge between

regions of differing A/B compartmentalization. Fit-Hi-C interactions can have both anchors interacting with compartment A or compartment B, or they can have one anchor interacting with each compartment.

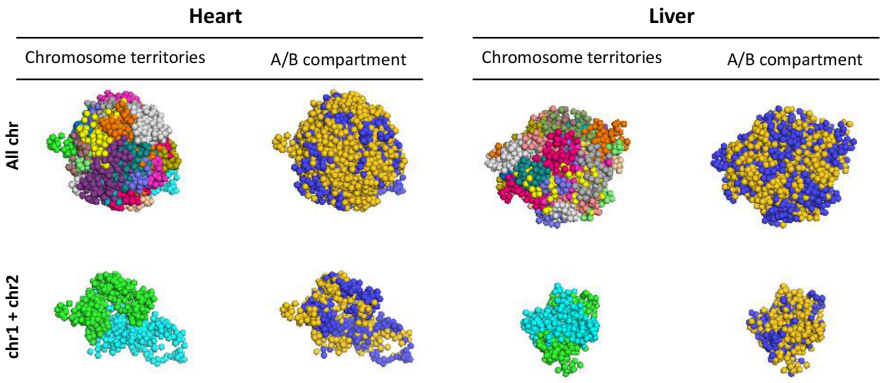


Figure 3-7. 3D models of liver and heart genomes. In both heart and liver nuclei, computational models generated using PGS reveal organization of topologically associating domains (TADs) within chromosomes in 3D space (top, All chr), as well as interactions between TADs of individual chromosomes (bottom, chr1 and chr2 shown as examples). Between chromosomes, regions of concordant A/B compartmentalization can aggregate in 3D space (blue = compartment A, gold = compartment B; See Supplemental Movie 1 for 360º view of this 3D reconstruction).

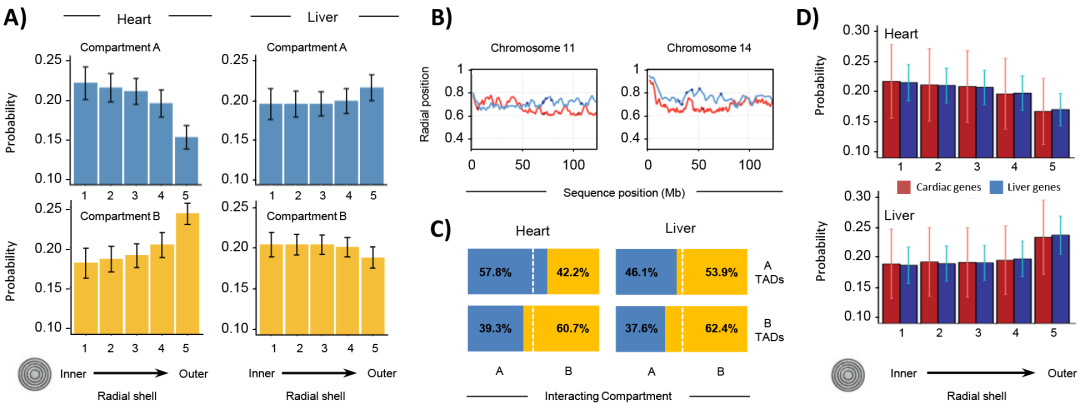


Figure 3-8. Properties of 3D models of heart and liver genomes. (A) The nucleus is divided into 5 concentric shells, with shell 1 in the interior and shell 5 at the periphery. In heart nuclei (left), TADs in compartment A are more likely to be found towards the interior of the nucleus and TADs in compartment B towards the periphery. In contrast, the liver Hi-C data (right) show that TADs in compartment A have a higher probability of being found towards the periphery and those in compartment B within one of the inner shells. Error bars indicate standard deviation of observations for 10,000 structures. (B) Across distinct regions of chr11 (left) and chr14 (right), radial positions of TADs between organs (heart in red, liver in blue) differ. Positions of heart- (red dots) and liver-specific genes (blue dots) superimposed. The y-axis shows average position (0 is the center of the

nucleus, 1 indicates nuclear periphery), while the x-axis shows chromosome position in megabases. (C) In 3D cardiac nuclear models (left), on average 57.8% of queried A-compartment TADs form interchromosomal interactions (within 500nm) with regions in compartment A, while 60.7% of queried B-compartment TADs form interchromosomal interactions with regions in compartment B. Contrastingly, liver models show that 53.9% of queried A-compartment TADs are within 500 nm of regions in compartment B, while 62.4% of queried B-compartment TADs interact with the same compartment in a different chromosome. (D) Similar analysis as in (A), but with TADs that have heart- (red bars) or liver-specific (blue bars) genes. In cardiac nuclei (left), both heart- and liver-specific genes tend to associate with the nuclear center, while in liver nuclei the trend is the opposite. Error bars indicate standard deviation of observations for 10,000 structures.

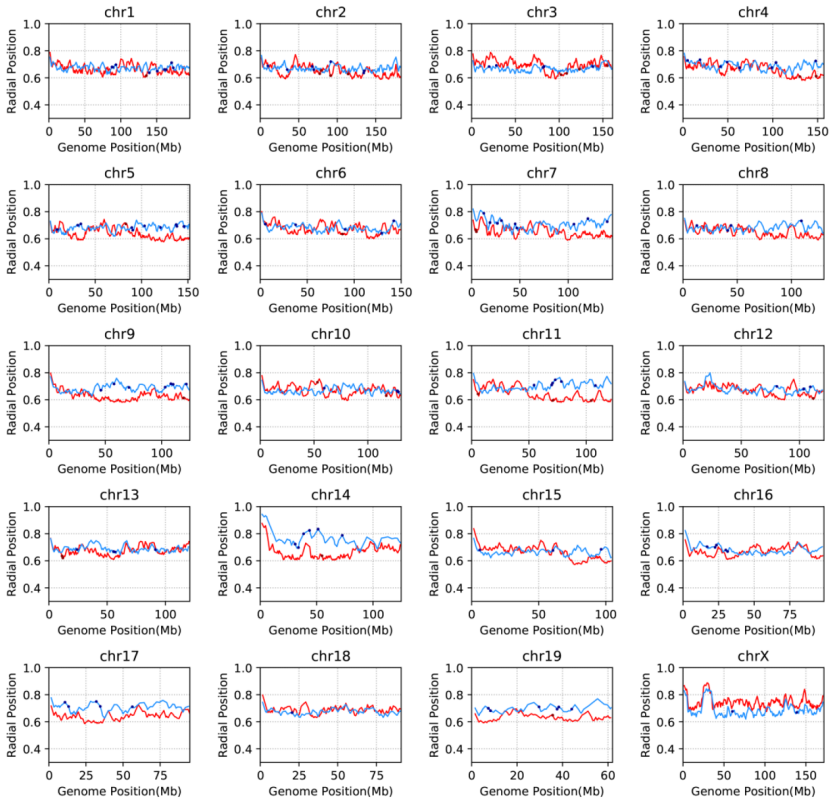


Figure 3-9. Radial positions of TADs differ between heart and liver. Across each chromosome, average radial positions of TADs are shown as solid lines (heart in red, liver in blue), with the positions of heart- (red points) and liver-specific genes (blue points) superimposed. The y-axis shows average radial position (0 is the center of the nucleus, 1 indicates nuclear periphery), while the x-axis shows the position of features along the chromosome.

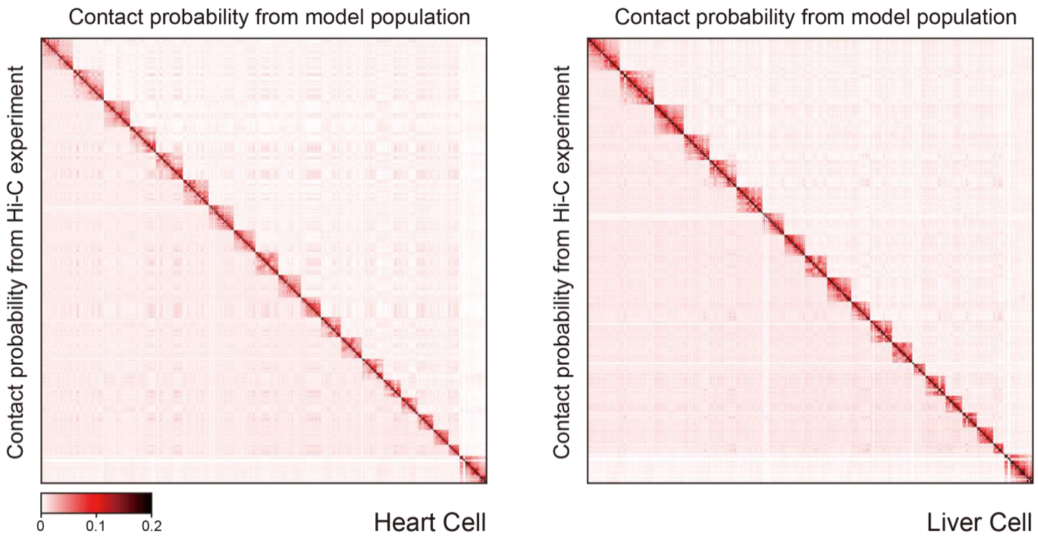


Figure 3-10. Comparison between contact probability heatmaps from experiment and structural models for heart (left) and liver (right). Each bin in the heatmap represents a TAD and each pixel represents the contact probability between 2 TADs. The lower triangle part shows the contact probability from experiment and the upper triangle shows the contact probability from the models. Contact patterns in the Hi-C experiment are very well reproduced in the structure models. The color scale ranges from 0 probability to 0.2 probability and any probability higher than 0.2 are shown as 0.2.

	Liver	Heart
Total_pairs_processed	1514557906	1362691412
Unmapped_pairs	21566557	11599265
Low_qual_pairs	0	0
Unique_paired_alignments	926874402	849101181
Multiple_pairs_alignments	408751544	385904274
Pairs_with_singleton	157365403	116086692
Low_qual_singleton	0	0
Unique_singleton_alignments	0	0
Multiple_singleton_alignments	0	0
Reported_pairs	926874402	849101181

	Liver	Heart
Valid_interaction_pairs	701407381	807707536
Valid_interaction_pairs_FF	167026398	200806229
Valid_interaction_pairs_RR	167121126	200745187
Valid_interaction_pairs_RF	167062459	199501877
Valid_interaction_pairs_FR	200197398	206654243
Dangling_end_pairs	140829780	20685150
Religation_pairs	80093214	17757504
Self_Cycle_pairs	2047729	967879
Single-end_pairs	0	0
Dumped_pairs	2496298	1983112

	Liver	Heart
Valid_interaction	701407381	807707536
Valid_interaction_rmdup	613559460	720072695
Trans_interaction	272280195	323763107
Cis_interaction	341279265	396309588
Cis_shortRange	83657373	74217618
Cis_longRange	257621892	322091970

	Liver	Heart
Fit-Hi-C interactions (q<0.01) INTRA	90587	115843
Fit-Hi-C interactions (q<0.01) INTER	94301	51370

Table 3-1. Summary of Hi-C data from heart and liver

Genes with Promoter-TES Interaction in Cardiac Hi-C Data					
Acads	Mb	Mydgf	Evx1os	Lao1	Zdhhc22
Adam12	Mbd3	Gnl3	Med26	Oas1d	Kcng2
Adcy6	Mbl1	Tor1a	Sik3	Nsun5	Pik3c2b
Aif1	Mbl2	Tor1b	Prss41	Upk3b	4933406M09Rik
Alox12e	Meis1	Slc46a2	Zfp689	Chpf2	D430041D05Rik
Aoc3	Mertk	Lsm4	Naif1	Snx21	Kcnk15
Ap2m1	Mip	Irx4	4933433G15Rik	Ric8a	Klhdc7a
Ap4m1	Meis2	Myadm	4933440J02Rik	Tubg1	Gzmn
Aqp3	Mtx1	Ccdc97	Treml1	Tubg2	Amdhd2
Arvcf	Myl3	Ldlrad4	Rnf215	Itga9	Ssh3
Bmp10	Myl2	Arhgef25	Tmem25	Cdc42ep2	Cd207
Bsn	Nab2	Dbndd2	Mus81	Slc39a3	Mfrp
C1qb	Ncf1	Ybx2	Pitpnc1	Brms1	Mpg
Capns1	Ndufv1	Prg3	1700003E16Rik	Mrpl10	Eme1
Cbl	Nfkb2	Rab25	Dpep3	Bre	Wdr24
Cct6a	Nfkbie	Ift20	Jsrp1	Shmt2	Dpcr1
Cd19	Nefm	Gtpbp2	Tmem79	Tbx10	Orai3
Cd37	Npdc1	Cpxm1	Ephx3	Hyal3	Ccdc153
Entpd1	Nthl1	Ppp4c	Gpc2	Slc35b1	Spib
Cd72	Nxn	Ctsf	Cdca4	Akr7a5	Fam205c
Ift81	Ogg1	Arl10	Erv3	Krt5	Pced1a
Cdx2	Olfr59	Apba3	Cyb5r1	Pigh	Gm20735
Cel	Pck1	Wdr46	Entpd8	Timp4	Ablim3
Celsr1	Pcsk4	Ppp1r1a	1810062G17Rik	Mbd6	3-Mar
Ckmt1	Pde6g	Cts6	Asb6	Clic1	Grin1os
Ccr2	Prf1	Rps6kb2	Klhl40	4930433N12Rik	Zfp629
Cnih2	Plau	Pnkp	2210016L21Rik	Cts8-ps	B230319C09Rik
Cnp	Plcb2	Slc22a17	Ccdc71	Cts3	B230206H07Rik
Col14a1	Serpinf2	Jph2	Aldh1b1	Gjc2	Mical2
Cpne6	Npy4r	Myg1	C1qtnf6	Grin3b	Trmt61a
Cryab	Mapk11	Fam129a	B3gat3	Znrf1	Lgr6
Cryba4	Klk6	Rbp7	Tymp	Socs7	C130026L21Rik
Cyp17a1	Cyth2	Slc29a1	2900057B20Rik	Rassf3	Zc3h4
Cyp1a1	Psmc3	Ctsm	Prss57	BC018473	Rnf150
Dapk2	Ptp4a3	Dpysl5	Tm7sf2	Zbtb7c	Krt78
Des	Igdcc3	Slc9a3r2	1700065J11Rik	Alg3	Morn2
Dlx4	Pxn	P3h4	Izumo1	Sirt7	Lrrc24
Ebf1	Rab20	Them5	1700067G17Rik	Eif2b1	A230065H16Rik
Eef2k	Rad9a	Nicn1	Mfsd12	Irgq	Prickle4
Efna1	Rai1	Pam16	Tm9sf1	Mrgprf	Stum
Egr3	Rara	Pnpla2	Uba7	Dcaf15	Sbk2
Ehd1	Rfxank	Tmx2	Perm1	Pgghg	Gm26705
Eif2b4	Rgl2	Plin5	2310065F04Rik	Crygn	Ccdc84
Eif4a1	S100a10	Tmem134	4931428F04Rik	Mmp21	Cdsn
Eif4ebp1	Ccl22	Cep19	4931430N09Rik	Rhot2	Thoc6
Eno3	Ccl9	Med29	4930478P22Rik	Slc43a2	C730036E19Rik
Fgf8	Nptn	Gpx7	Mamstr	Iba57	Dgkeos
Fhl4	Sh2b1	Qpctl	Nkx6-3	Spag7	D630024D03Rik
Fkbp10	Ptk6	4921536K21Rik	4930488B22Rik	BC030499	Itprip
Flt3l	Slc22a12	Ccdc127	Efcab10	Gm11545	Lyrm7os
Fosb	Slfn1	Map1lc3b	4930515B02Rik	Stxbp6	Trim80
Fut4-ps1	Sstr3	Abhd15	Dusp18	Gphb5	Ly6g6f
Slc37a4	Serpina3m	1700019N19Rik	Sirt4	Susd6	Jmjd7
Gata5	Srms	Nmral1	4930591E09Rik	Acot6	Gm13119
Gba	Aurkb	Cdca5	Slc6a21	Mfsd7c	Klhdc7b
Gcgr	Bhlhe40	Mri1	Tmem234	Rrp36	Cfap73
Gjc1	Stx4a	1110004E09Rik	Scarletltr	Apbb3	Gm15713
Gnat1	Tagln2	Lurap1	St5	Catsper1	D330041H03Rik
Rack1	Tcea2	Sdccag3	2510039O18Rik	Npas4	Gm7538
Grk5	Prdx2	Pomgnt1	Klhl2	Cfap157	Gm10433
Grb2	Tgfbr3	Crip2	9930111H07Rik	Large2	Gm16287
Hlx	Thy1	0610030E20Rik	Spata31	Trib3	D4Ertd617e
Nr4a1	Tlx1	Fam189b	Creb3l4	Hcrtr1	Gm13648
Hspa1l	Traf5	Wipf2	2210407C18Rik	Man1c1	Rpl37rt
Hsd11b2	Tuba1c	Syf2	Fam131a	Fbxo44	Gm10814
Hsp90ab1	Uck1	Fbxw9	B430010I23Rik	Plekhn1	Gm19510
Il11	Upk3a	Slc44a2	Cxcr6	Sds	Gm15408
Ilf3	Cdh23	Trp53inp2	Kat5	Fam109a	Gm13003
Ilk	Vav2	Acss1	Ankrd17	Mat2a	Cbarp
Irf1	Vegfb	Taf11	Wdr6	Arhgap35	Gm29805
Itgb5	Fmnl3	Gsdmd	Il24	Klc3	Gm12348
Kcna7	Xbp1	1700003D09Rik	Pcdhgc3	Tbc1d17	LOC105246506
Kcnab3	Ywhag	Tfpt	Arid1a	Prr14	
Kcnj12	Rnf112	2410004I01Rik	Cnnm2	Ccdc189	
Kcnk4	Zfp46	Nabp2	Tinagl1	Csgalnact1	
Krt33b	Zp3	Spaca9	Ptges2	Ces2b	
Krt1	Car14	Srrd	D930048N14Rik	Elmo3	
Krt84	Pappa2	Hk1os	Nol12	Carmil2	
Krt4	Impdh2	2010107E04Rik	Deptor	Cdk10	
Krt8	Mapk7	Plscr3	Phlpp1	6030466F02Rik	
Lat	Dazap2	Cactin	Ehd4	Atg4d	
Lhx3	B9d1	Reep6	Abtb2	Stac3	
Mycl	Naglu	Kptn	Tchh	Pla2g3	
Mafg	Nagpa	Dhrs13	Fam46b	Fscn2	

Table 3-2. List of genes with significant (q < 0.01) promoter-TES Fit-Hi-C interactions in the heart

Genes with Promoter-TES Interaction in Liver Hi-C Data			
Acr	Stat3	Efcab10	Gzmn
Aif1	Wnt8a	4930515B02Rik	Cd207
Alox12e	Sdc4	Spaca6	Slc5a2
Aqp2	Tfeb	Prr3	Olfr183
Aqp3	Tubb4a	4930555B11Rik	Olfr1414
Bsn	Ywhag	Ankrd53	Trim61
C1qb	Rnf112	Slitrk5	Krt12
C3	Zfp93	4930591E09Rik	E030025P04Rik
C4b	Pappa2	Rnf183	Mir124a-1hg
Cd24a	Fiz1	Fbxo31	Dpcr1
Cd37	Prnd	Got1l1	Zfp384
Cd82	Ing4	Slc6a21	Tm4sf19
Celsr1	Gnl3	St5	Klk15
Ccr1	Irx4	Fam162b	Klk14
Ccr5	Ech1	9930111H07Rik	A830011K09Rik
Col10a1	Lurap1l	B230112J18Rik	Ablim3
Cpne6	Prg3	Ccdc190	4732490B19Rik
Crx	Rab25	Tmem247	Tmem91
Csk	Chst2	Nol3	Il1bos
Cst7	Prss16	B430010I23Rik	Zfp78
Cyp1a1	Irgm2	Fbxo34	Serpina11
Cyp1a2	Il1f5	Dtx3	Gm1587
Des	Cts7	Siglece	Sbk3
Dhh	Cpxm1	Trem2	Sbk2
Dnase1l3	Dmrtb1	Pcdhgb8	Gm26705
Ecm1	Ick	Pcdhga12	Serpinb3d
Edn2	Kcne3	Pcdhb16	Krt87
Eno2	Asic5	Tchh	Lyrm7os
Ephx2	Ppp1r1a	Lao1	Ankrd13c
Fasl	Cts6	Oas1d	Gm1965
Fga	Rnf130	Upk3b	Tmem92
Flt3l	Gmfg	AI314278	Arsi
Fut4-ps1	Ctsm	Fam110c	Gpr17
Gjc1	Dpysl5	Gmpr2	9230009I02Rik
Gnl1	Gkn2	Mmrn2	Zfp551
Gna12	Ubxn6	Krt9	Gm11437
Gna15	Cep19	Rnf169	Angptl7
Got1	Ceacam11	R3hdm4	Gm16287
Lpcat3	Nde1	Pgc	D4Ertd617e
H2-Q4	Kcnv1	Krt5	Gm4719
H2-T10	Rnf41	Pigh	Krt83
Hpx	Polr2g	Timp4	Gm12505
Igfbp6	Sec14l2	Cd209d	Gm19345
Il12a	Nmral1	Cd209a	A730020E08Rik
Il2	Snrpg	Gpr37l1	Gm29805
Irx1	A930003A15Rik	Rassf3	Gm29669
Irx2	Rab13	BC018473	
Itih1	Tmem216	Pom121l2	
Kcna4	Nadk2	Dtx4	
Serpina3c	Lyzl6	Ghsr	
Krt19	2310034O05Rik	Zfp710	
Krt33b	Tatdn1	Best2	
Krt86	Prss32	Pth2r	
Krt84	Nabp2	Slc43a2	
Krt2	Spaca9	Slc36a1	
Mfap2	2610528A11Rik	Gm38403	
Fxyd3	Hk1os	Mgl2	
Cxcl9	2010107E04Rik	Trim25	
Myl3	Gfod2	Gm11545	
Myl2	Krt28	Mfsd7c	
Neurl1a	4933432G23Rik	Zfp503	
Nefm	Dmrtc2	Shisa2	
Nhlh2	4933430N04Rik	Rtp2	
Nov	4933433G15Rik	Speer2	
Nr4a2	Susd2	AI661453	
Nxn	Dpep3	Dpp9	
Pax4	Dnmbp	1110051M20Rik	
Pck1	Erv3	Gimap7	
Penk	Gcnt3	Lypd4	
Pomc	Spata31d1a	Acsm2	
Klk6	1700021F07Rik	Thumpd1	
Ptcra	1700030O20Rik	Ces2b	
Qk	Ces2g	Tat	
Rcvrn	Tagap	6030466F02Rik	
Resp18	Smc2os	Apeh	
Stmn3	Rassf6	Prss42	
Ccl6	Bpifa3	Pla2g3	
Sfrp2	1700067G17Rik	Spata31d1b	
Nptn	4930428G15Rik	Lrit1	
Slc31a1	Zswim5	Neu4	
Slfn1	Ifitm7	Kcng1	
Serpina3g	Mamstr	Cela3a	
Serpina3n	4930448H16Rik	Nyap1	
Spic	4930503B20Rik	Pskh1	

Table 3-3. List of genes with significant (q < 0.01) promoter-TES Fit-Hi-C interactions in the liver

Chapter 3: Bibliography

1. Stedman E, Stedman E. Cell specificity of histones. Nature 1950;166:780-1.

2. Allfrey VG, Faulkner R, Mirsky AE. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. Proc Natl Acad Sci U S A 1964;51:786-94.

3. Gold M, Hurwitz J, Anders M. The Enzymatic Methylation of Rna and DNA, Ii. On the Species Specificity of the Methylation Enzymes. Proc Natl Acad Sci U S A 1963;50:164-9.

4. Grunstein M. Nucleosomes: regulators of transcription. Trends in genetics : TIG 1990;6:395-400.

5. Strahl BD, Allis CD. The language of covalent histone modifications. Nature 2000;403:41-5.

6. Tremethick DJ. Higher-order structures of chromatin: the elusive 30 nm fiber. Cell 2007;128:651-4.

7. Lieberman-Aiden E, van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 2009;326:289-93.

8. van Steensel B, Dekker J. Genomics tools for unraveling chromosome architecture. Nature biotechnology 2010;28:1089-95.

9. Rao SS, Huntley MH, Durand NC, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 2014;159:1665-80.

10. Ay F, Bailey TL, Noble WS. Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. Genome Res 2014;24:999-1011.

11. Sayed D, He M, Yang Z, Lin L, Abdellatif M. Transcriptional Regulation Patterns Revealed by High Resolution Chromatin Immunoprecipitation during Cardiac Hypertrophy. J Biol Chem 2013;288:2546-58.

12. Dekker J, Mirny L. The 3D Genome as Moderator of Chromosomal Communication. Cell 2016;164:1110-21.

13. Schmitt AD, Hu M, Ren B. Genome-wide mapping and analysis of chromosome architecture. Nature reviews Molecular cell biology 2016;17:743-55.

14. Rosa-Garrido M, Chapski DJ, Schmitt AD, et al. High-Resolution Mapping of Chromatin Conformation in Cardiac Myocytes Reveals Structural Remodeling of the Epigenome in Heart Failure. Circulation 2017;136:1613-25.

15. Kim YH, Marhon SA, Zhang Y, Steger DJ, Won KJ, Lazar MA. Rev-erbalpha dynamically modulates chromatin looping to control circadian gene transcription. Science 2018;359:1274-7.

16. Servant N, Varoquaux N, Lajoie BR, et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol 2015;16:259.

17. Imakaev M, Fudenberg G, McCord RP, et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat Methods 2012;9:999-1003.

18. Shin H, Shi Y, Dai C, et al. TopDom: an efficient and deterministic method for identifying topological domains in genomes. Nucleic Acids Res 2016;44:e70.

19. Hua N, Tjong H, Shin H, Gong K, Zhou XJ, Alber F. Producing genome structure populations with the dynamic and automated PGS software. Nature protocols 2018;13:915-26.

20. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 2015;12:357-60.

21. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078-9.

22. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 2015;31:166-9.

23. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.

24. Lun AT, Perry M, Ing-Simmons E. Infrastructure for genomic interactions: Bioconductor classes for Hi-C, ChIA-PET and related experiments. F1000Res 2016;5:950.

25. Carlson M. KEGG.db: A set of annotation maps for KEGG. R package version 3.2.3. 2016.

26. Uhlen M, Fagerberg L, Hallstrom BM, et al. Proteomics. Tissue-based map of the human proteome. Science 2015;347:1260419.

27. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nature protocols 2009;4:1184-91.

28. Therizols P, Illingworth RS, Courilleau C, Boyle S, Wood AJ, Bickmore WA. Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. Science 2014;346:1238-42.

29. Reddy KL, Zullo JM, Bertolino E, Singh H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. Nature 2008;452:243-7.

30. Guelen L, Pagie L, Brasset E, et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature 2008;453:948-51.

31. Mathiyalagan P, Keating ST, Du XJ, El-Osta A. Chromatin modifications remodel cardiac gene expression. Cardiovasc Res 2014;103:7-16.

32. Haldar SM, McKinsey TA. BET-ting on chromatin-based therapeutics for heart failure. J Mol Cell Cardiol 2014;74:98-102.

33. Dickel DE, Visel A, Pennacchio LA. Functional anatomy of distant-acting mammalian enhancers. Philos Trans R Soc Lond B Biol Sci 2013;368:20120359.

34. Chang CP, Bruneau BG. Epigenetics and cardiovascular development. Annual review of physiology 2012;74:41-68.

35. Olson EN. A decade of discoveries in cardiac biology. Nature medicine 2004;10:467-74.

36. Movassagh M, Vujic A, Foo R. Genome-wide DNA methylation in human heart failure. Epigenomics 2011;3:103-9.

37. Sallam T, Sandhu J, Tontonoz P. Long Noncoding RNA Discovery in Cardiovascular Disease: Decoding Form to Function. Circ Res 2018;122:155-66.

38. Nothjunge S, Nuhrenberg TG, Gruning BA, et al. DNA methylation signatures follow preformed chromatin compartments in cardiac myocytes. Nature communications 2017;8:1667.

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

Future Directions: What's Next for Cardiovascular Epigenomics

The work summarized in this dissertation required years of experimental design and investigation, and established a foundation from which to generate hypotheses and advance the field of cardiovascular genomics. In addition, this work allowed our lab to develop and implement a plethora of bioinformatics and statistical frameworks to meet our basic research needs. We built our own server to perform big data analyses and wrote a large corpus of custom code to answer detailed biological questions. Our snapshot of how cardiovascular epigenomes change in heart failure provides rationale for us to ask *how* epigenomes change with time during pathological perturbation. Specifically, do epigenomic structural rearrangements from acute cardiovascular stimuli behave the same way as those resulting from chronic stimuli such as 3 weeks pressure overload or 5 weeks of tamoxifen-induced CTCF-knockout? I suggest a time course experiment, to understand the chromatin conformational *intermediates* that exist during the progression to heart failure, both from pressure overload and CTCF-knockout perspectives, accompanied by rigorous statistical tests to understand how chromatin structures change over time. Because heart failure can take years to develop in patients, this experiment would provide more mechanistic insight into how genome architectures change with heart disease, but also give us information about *when* they change.

Another avenue for cardiovascular chromatin structural research is to go to deeper resolution in the three-dimensional models that we make with Hi-C data[1]. This will allow us to integrate local topological changes that modulate gene expression along with higher-order chromatin dynamics to better understand how chromatin features change in three-dimensions from locus-to-nuclear scale, with relative positional information to drive

future experimental design. Such an endeavor would require physical computational infrastructure and further honing of tools to meet the bioinformatics demand of the question: First, we would need to deploy the mathematical modeling tools to a cloud-based resource such as Amazon Web Services, to afford us the computational scale and flexibility of job scheduling that might not be readily available on a university computing cluster. Second, to improve the quality of the analyses that occur after 3-dimensional modeling, modularization of the code for *ad hoc* analyses would be prioritized. This would make the generation of higher-resolution models faster (via cloud computing) and the exploration of biological questions more direct (via pre-written code). The codebase we have developed (and are still optimizing) will provide code accessibility to other researchers who are not as experienced with cardiovascular epigenomics, while maintaining rigorous reproducibility standards via resources such as Docker[2] to contain our analysis frameworks. A "Dockerized" workflow would allow us to deploy bioinformatics pipelines with a short learning curve and increased flexibility to provide internal and external collaborators with fast solutions at a lower monetary cost.

Another fundamental question with regard to basic mechanisms of pathology is whether different cell types behave the same way at the chromatin structural level with disease. Does a cardiac fibroblast epigenome, for instance, behave differently than that of a cardiac myocyte? The former undergoes cell division while the latter does not, but we hypothesize that the genomes for both cell types undergo chromatin architectural rearrangements since they are subjected to similar mechanical and signaling perturbations during disease[3]. The question then becomes how are these structural

changes conferring a cardiomyocyte- or fibroblast-specific phenotypic response downstream of cardiac stress?

From a translational side, we have unique tools for drug target identification: The hybrid mouse diversity panel (HMDP) provided rationale to study CTCF in the context of heart failure because the gene is downregulated in most mouse strains after long-term beta-adrenergic stimulus[4,5]. This resource will continue to drive future investigation into proteins that become dysregulated with disease. In addition to single-gene investigation, we can begin to understand the chromatin structural implications of genetic diversity from a computational perspective. For example, we can begin to ask whether modules of co-regulated genes in the HMDP interact with each other more than with other modules in 3-dimensional space (using chromatin conformation capture techniques followed by bioinformatics). Another translational implication of our work is the potential reversion of CTCF-knockout mediated cardiac phenotype if we put CTCF back into failing cardiac myocytes. Now that we have a robust cardiomyocyte-specific CTCF knockout model[4], we can explore what happens to cardiac muscle physiology when we restore CTCF levels in hearts that are already failing. Would adding back CTCF revert the structure back into its native/healthy configuration? If so, would the transcriptome begin to resemble that of a quiescent cardiac myocyte? A virus containing a CTCF construct could be used to restore CTCF levels in a mouse that has undergone cardiomyocyte-specific CTCF-knockout, and we could measure ejection fraction and left ventricular wall dimensions to determine whether a reversion to healthy phenotype is plausible with this gene therapy. Translationally, this is just the beginning of exciting work with epigenomes in the heart.

Future Directions: Bibliography

1. Chapski, D. J., Rosa-Garrido, M., Hua, N., Alber, F. & Vondriska, T. M. Spatial Principles of Chromatin Architecture Associated With Organ-Specific Gene Regulation. *Front. Cardiovasc. Med.* **5,** 186 (2018).
2. Boettiger, C. An introduction to Docker for reproducible research. *ACM SIGOPS Oper. Syst. Rev.* **49,** 71–79 (2015).
3. Saucerman, J. J., Tan, P. M., Buchholz, K. S., McCulloch, A. D. & Omens, J. H. Mechanical regulation of gene expression in cardiac myocytes and fibroblasts. *Nat. Rev. Cardiol.* (2019). doi:10.1038/s41569-019-0155-8
4. Rosa-Garrido, M. *et al.* High-Resolution Mapping of Chromatin Conformation in Cardiac Myocytes Reveals Structural Remodeling of the Epigenome in Heart Failure. *Circulation* **136,** 1613–1625 (2017).
5. Rau, C. D. *et al.* Mapping genetic contributions to cardiac pathology induced by beta-adrenergic stimulation in mice. *Circ. Cardiovasc. Genet.* **8,** 40–49 (2015).