Molecular mediators of cardiac-specific enhancer activation

by

Olukunle O. Demuren

A.B. Molecular Biology
Princeton University, 2011

SUBMITTED TO THE DEPARTMENT OF BIOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN BIOLOGY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SEPTEMBER 2020

Signature of Author: _____
Department of Biology
June 1st, 2020

Certified by: _____
Laurie A. Boyer
Professor of Biology and Professor of Biological Engineering
Thesis Supervisor

Accepted by: _____
Mary Gehring
Associate Professor of Biology
Member, Whitehead Institute
Co-Director, Biology Graduate Committee

Abstract

Understanding how transcription factors (TFs) control gene expression programs is critical for determining the genetic pathways responsible for development. Heart development is particularly sensitive to precise control of gene programs as faulty regulation leads to congenital heart defects (CHD), the leading cause of infant mortality. Although sets of TFs have known roles in heart development, in most cases, we lack a fundamental understanding of how these binding events regulate cell specification. To identify potential key regulatory TFs, we used the Assay for Transposase-Accessible Chromatin (ATAC-seq) to map changes in chromatin accessibility and integrated these data with maps of histone modification patterns and gene expression across several stages of embryonic stem cell (ESC) differentiation toward cardiomyocytes (CMs). Based on bioinformatic analysis of these data, we identified the TEA domain family (TEAD) TF TEAD1 as a candidate regulator of enhancer activation during cardiac-lineage commitment. We then used an inducible degron-tag strategy to conditionally deplete TEAD1 and observed an abnormal beating phenotype in CMs. Further mechanistic studies revealed that TEAD1 was necessary for the activity of a subset of cardiac enhancers putatively linked to cell-cell contacts. These data have allowed us to characterize a potential link between extracellular signaling and cardiac contraction and morphogenesis during development.

Acknowledgements

First of all, I want to thank my advisor, Laurie Boyer. I am so grateful to have had the opportunity to be in your lab and learn from your experience. I deeply appreciate your support throughout this extraordinary journey.

I also need to thank my thesis committee, Chris Burge and Mary Gehring. They always asked the right questions and encouraged me to keep my focus on solving the important problems. I learned a great deal from you both about communicating my science and working through challenges.

I would also like to thank other faculty in the Biology department whose advice about science and other matters was invaluable to me in completing this journey. Frank Solomon, Alan Grossman, and Steve Bell have all been immensely supportive to me at various times when I really needed it.

I would not have joined the Boyer Lab if not for the wonderful community that made me feel immensely welcome from the very beginning. Alex, Constantine, Vera, Pema, Kirsten, Meghan, and Jovi have all been incredibly helpful in helping me get over the finish line, and in making our lab a great place to come to work. In terms of lab alumni, I want to specifically thank Joe, who was a very patient and accessible mentor during our time working together; Lucy, who was the best colony-picking UROP any graduate student could ask for; Gizem, for putting up with my mess when we shared a bay; and Darlene, whose wit, generosity, and organization made my life so much easier. I also have to thank Lauren, Sera, Vidya, Xinchen, Paul, Arune and Lilly for their friendship during their time in the lab and long after.

The second floor of Building 68 has been a great place to do science for the past eight years. All of the labs on the floor have been immensely helpful with reagents, informative discussions, advice, and camaraderie, but I must specifically thank our original next-door neighbors, the Saeij Lab, for all the joyful times and many Thanks-O'Ween celebrations, and the Burge Lab, who practically adopted me as a member when I really needed a supportive community.

I would be remiss if I didn't make sure to acknowledge the people who helped get me started in my path in biological research. The SRPG at the Broad Institute was just what I needed to get a taste of what might be possible in science – thanks to Rob Riley for taking me on for that summer, and my deepest appreciation for Bruce Birren, Eboney Hearn, and Francie Latour, whose support and mentorship during my time at the Broad and beyond has been so valuable. Hilary Coller was my undergraduate thesis advisor at Princeton, and she is still one of the kindest people I know in science.

My classmates have also been a great group to come through this program with. I will always remember Pit Parties, planning the Halloween party, and so many other good times. I was also lucky enough to come to MIT with three of my Integrated Science classmates from Princeton, completing a long scientific journey that started in the whale with the basketball problem – it has been a great pleasure to go through this with Mike, Aaron, and Mina.

I was also lucky to join an amazing community of friends through living at the fabled house at 65 Josephine Avenue. Grace, I look forward to watching many more action blockbusters with you. In particular, though, Sean and Alex have been putting up with me from Josephine to Putnam. Thank you both for all the advice, commiserations, jokes, and support. My life has been incalculably enriched by living with you two.

I also have to thank some of my closest friends who helped me maintain my link to a world outside of the lab. Thank you to the denizens of Rundleland, whose rowdy discussions about trivia and other things have helped keep me sane. Thanks also to my group of Princeton friends whom I have built so many memories with even after we left the Orange Bubble, at weddings, Reunions, fantasy football drafts and the like. I've learned that sometimes, you're just the goose. Specifically, I have to thank Nida, whose generosity of time and spirit knows no bounds and whose insight has taught me so much; Amanda, whose spirit continually inspires me; and Chantel, who never lets me take myself too seriously.

Finally, I have to thank my family. I have been so lucky to be close to my brother Ayotunde and my sister Olutoyin while they were at Harvard, and I am so proud of the people they have become. Ayotunde, thank you for always being there to talk soccer with me and share a joke when I needed a break. Olutoyin, thank you for being a source of wisdom beyond your years. Most of all, though, I have to thank my parents. My dad, on top of being the first reader of this thesis, has been so helpful in pushing me on and letting me learn from his own experience, and my mom has always been there with support and encouragement, and making sure that I was always taking care of myself. Together, they raised me to be an inquisitive and determined person, and have done so much to provide me with every opportunity possible. My parents left Nigeria over thirty years ago to start a new life in this country, and this accomplishment is just one of the fruits of their labor and sacrifice.

Table of Contents

Table of Figures

Chapter 1

Chapter 2

Chapter 1

Introduction

Author Contributions for Chapter 1:
Adapted by Olukunle Demuren in part from a review article written by Joseph Wamstad, Xinchen Wang, Olukunle Demuren, and Laurie Boyer; figures by Olukunle Demuren.

Wamstad, J.A., Wang, X.*, Demuren, O.O.*, and Boyer, L.A. (2014). Distal enhancers: new insights into heart development and disease. Trends in Cell Biology *24*, 294–302. *equal contributors

Significance

Understanding the mechanisms that underpin the precise regulation of gene expression during mammalian development remains a fundamental biological question. Proper control of developmental programs in the heart is particularly evident during lineage commitment, when complex gene expression patterns are exquisitely regulated during differentiation and morphogenesis (O'Meara et al., 2015; Wamstad et al., 2012). Consequently, even subtle perturbation of transcriptional networks leads to congenital heart defects (CHD) and some forms of adult heart disease (Bruneau, 2008; McFadden and Olson, 2002). The genetic factors that contribute to heart development have been studied in a range of organisms such as zebrafish, fruit flies, mice, and humans (Bruneau, 2013; Olson, 2006); however, the majority of genetic causes of CHD remain unknown (Blue et al., 2017). Decades of research have led to the identification of integrated gene regulatory networks comprising a critical set of cardiac transcription factors (TFs) that coordinate embryonic heart development (Akerberg et al., 2019; Olson, 2006; Spitz and Furlong, 2012; Srivastava, 2006). Chromatin-remodeling and modifying factors also have important roles in regulating cardiovascular development (Chang and Bruneau, 2012). In fact, many mutations linked to CHD are found within genes encoding TFs and chromatin regulators (Homsy et al., 2015; Jin et al., 2017; Zaidi et al., 2013). Together, lineage-specific TFs and chromatin regulators coordinate the activation and repression of complex transcriptional networks and emerging evidence indicates that selective binding of these factors to enhancer elements is a key part of this process. Our prior work has shown that the enhancer landscape is highly dynamic during heart development, and activation of these non-coding DNA elements underlies the programmed shifts in cell identity that lead to a fully formed and functional heart (Wamstad et al., 2012, 2014). The work in this thesis extends these findings and focuses on identifying the key mechanisms by which specific TFs coordinate gene expression programs through tissue-specific enhancer activation. Results from these studies will contribute critical pieces of the puzzle for understanding the genes and processes that orchestrate heart development and how faulty regulation can lead to CHD.

Enhancers regulate tissue-specific gene expression patterns

Enhancers are *cis*-regulatory elements that can activate the expression of distal target genes. These elements are bound by a constellation of lineage- and signal-dependent TFs that recruit co-factors such as chromatin modifiers and remodelers, setting off a cascade of events that culminates in the assembly of the transcriptional machinery at target promoters. The idea that enhancer elements contribute to gene activation was first illustrated over thirty-five years ago based on pioneering work using SV40 viral enhancers to drive transcription *in vitro* (Banerji et al., 1981; Moreau et al., 1981). Enhancer trapping experiments in *Drosophila*, in which an exogenous reporter is introduced into the genome via a transposable element, provided further critical evidence that distal non-coding DNA sequences could function to mediate precise spatial and temporal control of gene expression during metazoan development (O'Kane and Gehring, 1987). Enhancers have subsequently been shown to function at variable distances and independent of orientation from the genes they regulate to promote expression of reporter genes both *in vitro* and *in vivo* (Atchison, 1988; Bulger and Groudine, 2011; Maston et al., 2006). Given that enhancers generally lack a consensus DNA sequence, it had previously been difficult to identify sets of enhancers for any given cell type.

Cell-type-specific TFs bind disproportionately within putative enhancer regions, suggesting that the co-occupancy of different TFs offers the necessary versatility to direct cell-type-specific gene expression. Consequently, the binding of clusters of these TFs has provided a means to identify enhancer elements using high-throughput sequencing (Akerberg et al., 2019; Wamstad et al., 2014). The observation that histone post-translation modifications, incorporation of histone variants, and overall nucleosome positions differ at enhancers as compared to promoters (Andersson and Sandelin, 2019) suggests that enhancers have evolved distinct properties. Current models suggest that TF binding at enhancers transmits key signals to direct the spatiotemporal activation of developmental gene programs in response to cellular cues.

High-throughput analysis and prediction of enhancer elements

Technological advances have enabled genome-wide epigenetic profiling, providing global methods for defining putative enhancer elements. High-throughput assays used to identify potential enhancers rely on genome-wide measurements of epigenetic patterns that correlate with *bona fide* active enhancer regions, in contrast to low-throughput methods (Gasperini et al., 2020). The ability to globally identify putative enhancers has provided critical insights into developmental gene regulation, but the limitations of these techniques should be taken into account. Most hallmarks of enhancer activity are based on correlations, and experimental validation of these regions *in vivo* remains a key issue in the field (Halfon, 2019); as such, new methods that robustly identify *bona fide* enhancers are needed to reduce false positives. For example, a recent study demonstrated that high-confidence enhancer predictions come from combining relatively independent hallmarks of enhancer activity and function (Akerberg et al., 2019). Thus, studies of enhancer-mediated gene expression should leverage multiple techniques for enhancer identification to ensure that insights have the strongest empirical basis.

Chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-seq) and related techniques allow genome-wide profiling of post-translational histone modifications and TF binding with the use of specific antibodies. Based on analysis of histone modification patterns, several reports suggested that mono-methylation of histone H3 lysine 4 (H3K4me1) marks a broad set of putative enhancers, whereas aceylation of histone H3 lysine 27 (H3K27ac) is associated specifically active enhancers (Creyghton et al., 2010; Ernst et al., 2011; Rada-Iglesias et al., 2010). In the absence of active enhancer-associated marks H3K4me1 deposited by the methyltransferases MLL3 and MLL4 (Herz et al., 2012) marks poised enhancers, defined by the potential to become active in later developmental lineages. On the other hand, the deposition of H3K27ac in distal regions, which is catalyzed by the histone acetyltransferase (HAT) p300, strongly correlates with TF binding and transcriptional activation of target genes. Notably, p300 appears to bind significantly fewer enhancers than those identified by H3K27ac, perhaps suggesting that the modification represents a more stable mark of p300 activity.

Putative enhancers identified by histone modification patterns are highly unique in each cell type and neighbor genes with cell-type-specific functions (Barski et al., 2007; Bonn et al., 2012; Ernst et al., 2011; Rada-Iglesias et al., 2012; Visel et al., 2009; Wamstad et al., 2012). With the inclusion of genomic binding patterns of TFs and chromatin regulators as well as patterns of accessible chromatin and distal transcription, hundreds of thousands of putative enhancers in metazoans have been identified across different cell types and tissues in flies, mice, humans, and other (Akerberg et al., 2019; Andersson et al., 2014; Birney et al., 2007; Consortium et al., 2014; Dunham et al., 2012; Kharchenko et al., 2010; Kundaje et al., 2015; Shen et al., 2012; Thurman et al., 2012; Visel et al., 2007). The ability to identify putative enhancers at global scale has opened the door for understanding development, organismal complexity, and evolution.

The application of various techniques using *in vivo*-derived heart samples has led to the identification and validation of enhancers in heart development that regulate the expression of important cardiac genes (Akerberg et al., 2019; Dickel et al., 2016; May et al., 2011); however, these studies did not examine the dynamics of lineage commitment. To this end, our lab adapted an *in vitro* cardiomyocyte (CM) differentiation system using mouse ESCs that progresses through four stages (ESC, mesoderm (MES), cardiac progenitors (CP), and CM) by adding growth factors and cytokines at defined time points (**Figure 1A**) (Wamstad et al., 2012). We found that H3K4me1 and H3K27ac mark candidate poised and active enhancer regions (**Figure 1B**), identifying over 80,000 putative enhancers across the developmental time course (**Figure 1C**). Notably, putative active enhancers displayed very low overlap between closely related cell types and were highly unique to each stage, illustrating the rapid transitions of chromatin states in regulatory regions as cells differentiate. This study also predicted TFs that orchestrate heart development by combining motif enrichment within active enhancers and gene expression of the corresponding TF family members. Our work revealed enrichment of TFs with known roles in heart development and also suggested unexpected linkages between TFs at cardiac enhancers. For example, the motifs for MEIS1 and GATA4, two TFs not previously known to act in concert during heart development, co-occur in many of the same enhancers

(**Figure 1D**). In enhancer reporter assays, these factors worked together to increase gene expression. Collectively, our study demonstrated the value of high-throughput enhancer predictions for deriving insights into developmental gene control.

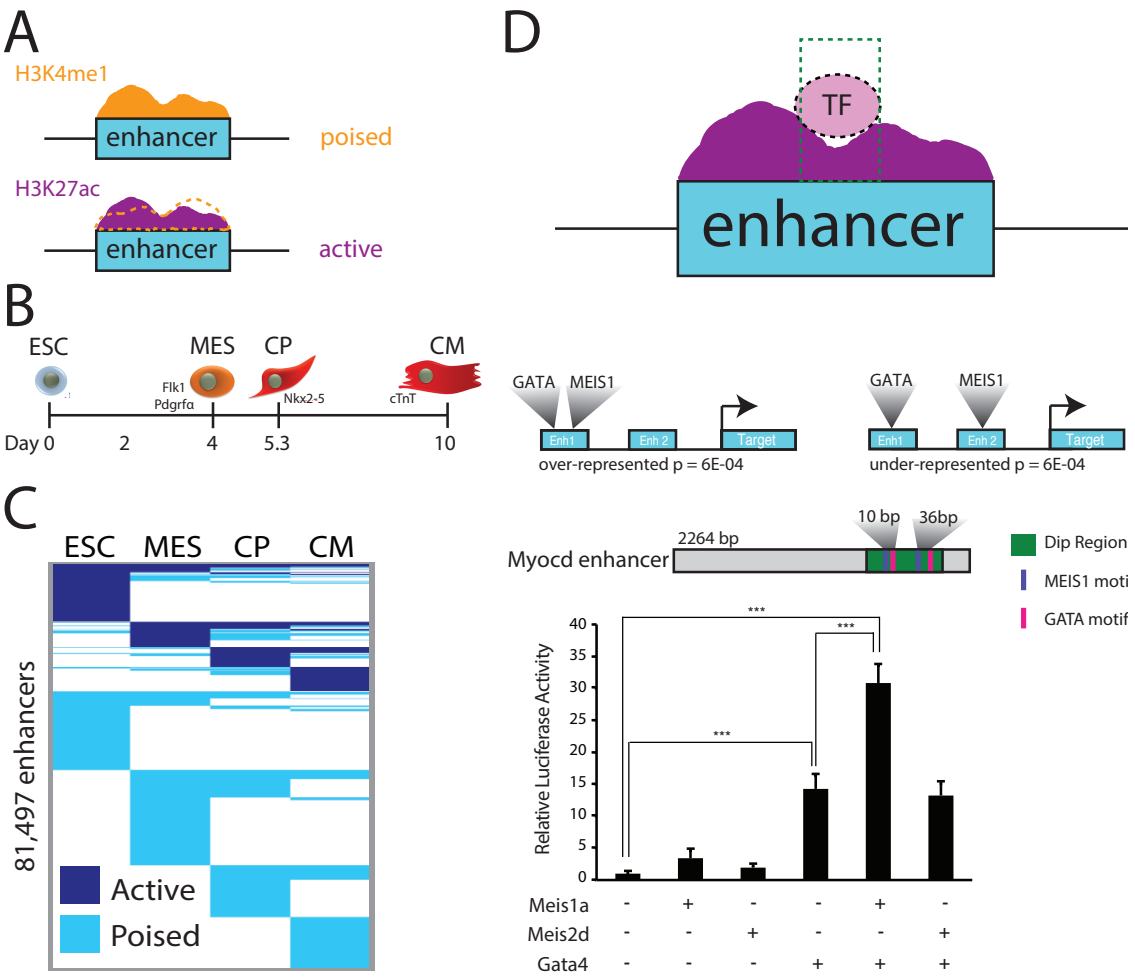


Figure 1: Enhancer Landscape Dynamics During CM Differentiation

A) Depiction of enhancer annotations based on histone modification patterns. Putative enhancers that are marked with H3K4me1 only are annotated as poised, and enhancers marked with H3K27ac with or without H3K4me1 (indicated by the dotted line) are annotated as active.

B) Schematic of *in vitro* cardiomyocyte (CM) differentiation. Embryonic stem cells (ESC) progress to pre-cardiac mesoderm (MES) in an embryoid body stage, and then differentiate to cardiac progenitors (CP) and cardiomyocytes (CM) in a monolayer upon treatment with the cytokines and growth factors indicated at top in blue. Factors used to assay the quality of differentiation are indicated in black.

C) Heatmap depicting dynamics of enhancer states during *in vitro* CM differentiation. Dark blue indicates active enhancers, light blue indicates poised enhancers, and white indicates unmarked enhancers. Adapted from Wamstad et al. (2012).

D) Identification of new TF partnership in CM differentiation. Dips (green dotted box) in H3K27ac signal represent potential TF binding sites. These sequences were used as input for motif analysis. The GATA and MEIS motifs were enriched within dips in the same enhancer more often than in separate enhancers, and a synthetic reporter assay shows that these factors synergistically activate an enhancer. Adapted from Wamstad et al. (2012).

Other studies have shown that the co-occupancy of cardiac TFs can be used to identify cardiac-specific enhancers (Akerberg et al., 2019; He et al., 2011; Junion et al., 2012). By performing ChIP-seq in HL-1 murine CM cell lines engineered to overexpress biotinylated isoforms of the key cardiac transcription factors GATA4, MEF2A, TBX5, and NKX2-5, and combining those data with p300 ChIP-seq, investigators identified thousands of putative enhancers co-bound by these factors. Importantly, roughly half of the putative enhancers tested drove cardiac-specific expression of reporter constructs in mouse embryos (He et al., 2011). A follow-up to this study used a similar approach to survey the binding of a wider set of cardiac TFs (adding TEAD1, SRF, and MEF2C) in fetal and adult murine hearts (Akerberg et al., 2019). This work found that even in the absence of H3K27ac, the binding of a large proportion of cardiac TFs could identify *bona fide* enhancers.

Evidence for combinatorial interactions between TFs at cardiac enhancers has also been demonstrated in *Drosophila* where the collective binding of five cardiac TFs was required for full transcriptional activity of heart enhancers *in vitro* (Junion et al., 2012). The approach of using genome-wide histone modification data to identify cardiac-specific enhancers, combined with analysis of genomic sequences to uncover regulatory TF networks has also been used to strong effect in *Drosophila*. Ahmad et al. analyzed data on verified enhancers from heart muscle cells, non-muscle heart cells, and pre-cardiac mesoderm using a machine learning framework to both predict other potential cell-type-specific enhancers and identify the particular motifs and corresponding TFs important for these classifications (Ahmad et al., 2014). Another study used a similar approach for enhancers active late in cardiogenesis (Seyres et al., 2016). Using solely a computational approach, other work identified over 40,000 putative cardiac enhancers in the human genome based on various sequence features including the co-occurrence of TF motifs (Narlikar et al., 2010). Remarkably, 62% of the candidate enhancers tested (16/26) in that study recapitulated cardiac gene expression in zebrafish reporter assays, indicating that motif composition may also provide a robust method for predicting functional enhancers. Taken together, these results show that a relatively small subset of TFs work together through their

binding at enhancers to control tissue-specific gene expression. Despite this growing knowledge, how TF partnerships at enhancers might determine specific aspects of heart development is a major open question.

High-throughput sequencing has also facilitated the development of function-based screens for identifying enhancer elements. STARR-seq (self-transcribing active regulatory region sequencing) uses random genomic DNA fragments to identify genomic sequences with the ability to activate gene expression in *Drosophila* (Arnold et al., 2013a; Yáñez-Cuna et al., 2014). STARR-seq is an extension of the MPRA (massively parallel reporter assay), a protocol based on saturation mutagenesis of specific enhancers in order to dissect function base-by-base (Melnikov et al., 2012; Patwardhan et al., 2012). One caveat of these approaches is that by removing sequences from their endogenous context, some information is lost about the precise context in which these loci work to activate gene expression. Nevertheless, these techniques offer direct assessment of the ability of candidate sequences to activate downstream transcription in an unbiased manner. As a complement to these approaches, the emergence of the CRISPR/Cas9 system has facilitated the testing of endogenous enhancers in a high-throughput manner (Klein et al., 2017). Various studies have combined scanning guide RNAs across loci of interest with relevant functional readouts to assess the importance of specific enhancer sequences in controlling target gene expression (Canver et al., 2015; Korkmaz et al., 2016; Sanjana et al., 2016).

Systems such as CRISPR/Cas9 or the lower-throughput TALEN have recently been applied to modulate the epigenetic state of enhancers. For example, the TALEN system was used to target the histone demethylase LSD1 to silence specific enhancers (Mendenhall et al., 2013). Further development of a catalytically inactive Cas9 (dCas9) (Qi et al., 2013) has allowed the targeting of epigenetic effectors to test the functionality of specific enhancers. For example, one study used dCas9 tethered to the repressive KRAB domain or to the p300 activator in order to target enhancer elements surrounding either the HER2 or the beta-globin loci in human cells (Klann et al., 2017). This study identified previously unknown enhancer regions, as well as showed that

some of the enhancers have cell-type-specific activity, providing a high-throughput assessment of enhancer activity in an endogenous context. This method of targeting epigenetic activators to putative enhancers has also identified key regulatory regions by their ability to increase target gene expression (Hilton et al., 2015; Simeonov et al., 2017). Many of these studies use fluorescence-based reporter readouts, which are more generalizable for a broader range of genes, but requires further genetic modification (Diao et al., 2016; Klann et al., 2017; Rajagopal et al., 2016). Other studies in this vein have relied on readouts in proliferation (Fulco et al., 2016), oncogene-induced senescence (Korkmaz et al., 2016), or resistance to a cancer treatment (Sanjana et al., 2016) that are correlated with the expression of a target gene.

Recent studies have leveraged advances made in single-cell gene expression technology in order to substantially increase the throughput of CRISPR-based enhancer perturbation screens. Single-cell RNA-seq was used to assess enhancer perturbation (dCas9-KRAB repression) by analysis of gene expression changes, showing that among a cell population, different enhancers within a super-enhancer cluster (defined below) can contribute differently to gene expression (Xie et al., 2017). Using an approach inspired by expression quantitative trait locus (eQTL) studies, Gasperini et al. developed a method for measuring the impact of enhancer perturbation on gene expression by lentiviral transduction of a gRNA library at relatively high multiplicity-of-infection (MOI) into dCas9-KRAB-expressing cells (Gasperini et al., 2019). This approach allowed high-throughput functional assessment of enhancer function by comparing gene expression in cells with and without repression of the targeted enhancer.

The above approaches depend on measurable changes in gene expression. Moreover, using a repressor such as KRAB or an activator like p300 relies on the underlying assumption is that modifying the chromatin state of an enhancer will have direct consequences on its ability to activate gene expression. Thus, these attempts to study enhancers in their endogenous contexts may still be missing important insights. Despite these caveats, advancements in high-throughput sequencing and the ability to target specific enhancers by genomic and epigenomic

perturbation have enabled the identification and functional characterization of putative enhancers at scale.

Predicting enhancer target genes

The characterization of putative enhancers in multiple cell types has led to the development of a range of techniques for the systematic identification of their target genes (**Table 1**). Advances in techniques such as Hi-C, which measure spatial proximity between genomic sequences in a multiplex fashion, have made it possible to characterize enhancer-target gene interactions genome-wide (Dixon et al., 2015; Jin et al., 2013; Lieberman-Aiden et al., 2009). Hi-C has been a valuable tool to study how enhancers direct developmental gene expression (Bonev et al., 2017; Phanstiel et al., 2017). Because Hi-C captures many interactions beyond those between enhancers and target genes, further developments have amplified the signal of interactions of interest. HiChIP (coupling Hi-C with chromatin immunoprecipitation) with H3K27ac antibodies can link active enhancers with the genes that they are putatively regulating (Mumbach et al., 2017; Weintraub et al., 2017), as can ChIA-PET (Chromatin Interaction Analysis with Paired-End-Tag sequencing) with RNA polymerase II (RNAPII) antibodies (Fullwood et al., 2009; Li et al., 2012). Other modifications of the original Hi-C method limit interactions to promoters. Capture Hi-C (also known as promoter capture Hi-C) as well as the similar HiCap use biotinylated nucleic acid probes to select for interactions with promoter sequences, increasing the signal of promoter-based interactions (Mifsud et al., 2015; Sahlén et al., 2015). Capture Hi-C has been used to reveal enhancer dynamics during developmental progression (Freire-Pritchett et al., 2017; Kragesteen et al., 2018; Rubin et al., 2017). This technique has also been invaluable for linking single-nucleotide polymorphisms (SNPs) associated with cardiovascular disease with potential disease genes by connecting enhancers containing these SNPs to target genes (Montefiori et al., 2018).Collectively, These protocols have greatly increased our ability to directly link enhancers to the genes they regulate, facilitating further mechanistic insights.

	method	description	advantages	limitations	references
direct measurement	Hi-C	all-to-all proximity ligation with paired-end sequencing	direct measurement of long-range interactions	low resolution, requires many reads	Liebeman-Aiden et al. (2009)
	Capture Hi-C	Hi-C (6-cutter) with promoter capture using biotinylated RNA probes	selects for promoter-linked interactions	limited to interactions involving probed regions	Mifsud et al. (2015)
	HiCap	Hi-C (4-cutter) with promoter capture using biotinylated RNA probes	higher resolution than CHi-C due to shorter fragments	harder to capture longer-range contacts	Sahlén et al. (2015)
	ChIA-PET	ChIP followed by proximity ligation and paired-end sequencing	enrichment of meaningful interactions	high number of cells needed	Fullwood et al. (2009), Li et al. (2012)
	HiChIP	Hi-C with antibodies to capture contacts containing protein of interest	lower cell number and higher resolution than ChIA-PET	limited by available antibodies	Mumbach et al. (2016)
computational	correlation-based	measure correlation between enhancer activity and promoter activity/gene expression across multiple cell types	captures multiple enhancers/gene, high resolution, quantitative measure	data availability, poor predictive ability, doesn't consider cooperativity	Corradin et al. (2013), Shen et al. (2012), Thurman et al. (2012), Yao et al. (2015)
	machine learning	link enhancers and promoters through patterns of sequence features/epigenetics based on true positives and negatives	high resolution, determines matches in an unbiased manner, can make predictions in other cell types	depends on quality of true positive and negative matches, data availability	Hafez et al. (2017), He et al. (2014), Whalen et al. (2016)
	regression-based	determines contribution of each enhancer's activity to expression of target genes	combinatorial assessment, quantifies most important enhancers, can be combined with other approaches	computationally intensive for large numbers of enhancers, data availability	Hait et al. (2018), Cao et al. (2017), Roy et al. (2015)
	score-based	assigns a single score to each enhancer-gene linkage	relative ease of computation, quantifies enhancer-promoter linkages, can be tuned transparently	arbitrary weights assigned to different data, data availability	Fishlevich et al. (2017), Naville et al. (2015), Zhu et al. (2016)

Table 1: Methods for Identifying Enhancer Target Genes

Although these methods are currently considered gold standard approaches, these techniques remain technically challenging and the output data are often difficult to compare across multiple cell types. To complement these efforts, computational approaches have evolved in parallel to identify enhancer-gene linkages (Hariprakash and Ferrari, 2019; Lim et al., 2018). The simplest algorithms identify correlations between levels of enhancer activity and promoter or gene activity across cell types (Corradin et al., 2013; Shen et al., 2012; Thurman et al., 2012; Yao et al., 2015). These methods require consistent datasets from many cell types to be effective, and are thus susceptible to bias based on the available genomic data. However, they do allow a quantitative measure of enhancer-gene association and can identify interactions over relatively short genomic distances that many direct measurement techniques cannot detect.

More sophisticated approaches utilize machine learning with a training set of validated true positives and negatives (Hafez et al., 2017; He et al., 2014; Whalen et al., 2016). These algorithms allow for enhancer predictions in other cell types, but the need for reliable information on true positive and negative linkages between enhancers and target genes still limits these predictions. Regression-based approaches are designed to measure the

contribution of the activity of each enhancer to the expression of putative target genes, allowing identification of the most important enhancers to each gene (Hait et al., 2018) and can be combined with other approaches (Cao et al., 2017; Roy et al., 2015). Regression-based analyses also require a limit on the number of enhancers surveyed per gene due to computational complexity, which may bias the identification of relevant linkages. Finally, broader approaches assign a score to each pairwise enhancer-gene interaction based on a range of relevant characteristics like co-expression of bound TFs and target genes, genomic distance, similarity in chromatin state, and/or evolutionary conservation (Fishilevich et al., 2017; Naville et al., 2015; Zhu et al., 2016). Although the assessment of these scores is only limited by having the necessary data in the cell type of interest, the integration of disparate datasets requires careful scaling and assumptions about the importance of particular information. These algorithms provide a means to integrate the massive genomic data now available to make testable predictions about the relationship between cis-regulatory elements and gene networks. Overall, experimental and computational advances in genome-wide approaches have led to important insights into enhancer biology, which can be leveraged to understand dynamic cellular processes such as development.

Enhancers control gene expression through a variety of mechanisms

The precise activation of cell-type-specific gene programs involves enhancer and promoter loci brought into spatial proximity (Bulger and Groudine, 2010; Wamstad et al., 2014). Enhancers can directly regulate genes as far as 1 Mb away (Lettice et al., 2003), suggesting an active process for enhancer selection. Many factors have been implicated in the formation and maintenance of these long-range interactions. As discussed above, epigenetic events such as the deposition of specific histone modifications, increased chromatin accessibility, and the binding of cell-type-specific TFs all strongly correlate with enhancer activation. How these events are coordinated with activation of target gene expression still requires further study.

Since enhancer target genes are often expressed in a cell-type-specific manner, particularly during development, understanding the mechanisms that drive long-range interactions

between enhancers and their targets are key to understanding gene regulation. To this end, different models have been proposed to explain the nature of enhancer-promoter interactions during development (**Figure 2**). In the "permissive" model, enhancer-promoter interactions are established prior to gene activation earlier in development, allowing cell-type-specific TFs to rapidly activate gene expression in response to developmental cues (Laat and Duboule, 2013). In contrast, the "instructive" model posits that enhancer-promoter interactions are formed *de novo* as differentiation progresses, depending on the binding of lineage-specific TFs that promote loop formation by recruitment of co-activators and the transcriptional machinery (Laat and Duboule, 2013). The development of techniques for assaying long-range genome interactions has begun to shed light on regulation of these dynamics. From the initial development of Chromosome Conformation Capture (3C) that allowed detection of long-range interactions between pairs of genomic loci (Dekker et al., 2002), to recent advancements like 5C, Hi-C, and Capture Hi-C (discussed above) that allow more comprehensive genome-wide assessment of long-range interactions (Schmitt et al., 2016), we are now in a position to investigate the relative contribution of these models to the spatiotemporal regulation of developmental gene expression.

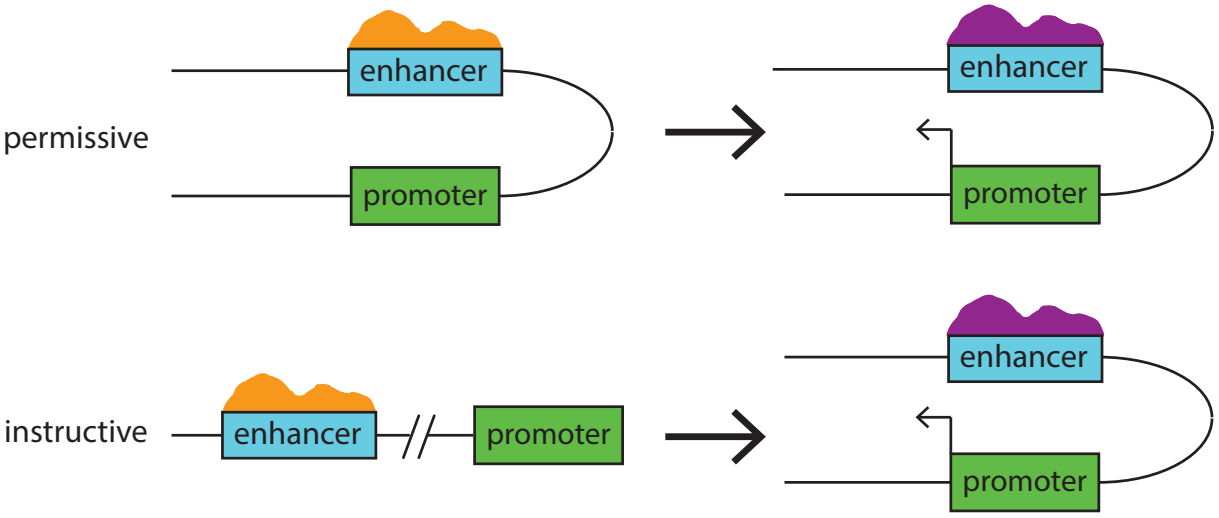


Figure 2: Models of the Establishment of Enhancer-Promoter Interactions
In the permissive model, enhancer-promoter interactions are established prior to gene activation, whereas the instructive model posits that loop formation is a driver of gene activation. Enhancers are marked prior to activation by the poised enhancer mark H3K4me1 (gold), and then marked by the active enhancer mark H3K27ac (purple). Gene activation is indicated by the arrow at the TSS.

Hi-C experiments have provided further insights into the nature of long-range enhancer interactions across a range of cell types (Bonev et al., 2017; Dixon et al., 2012, 2015; Jin et al., 2013; Lieberman-Aiden et al., 2009; Rao et al., 2014). These studies revealed that much of the topological structure of the genome is invariant between different cell types, particularly at the scale of topologically associating domains (TADs) (also known as contact domains). However, detailed studies of genome topology during development have shown that some of these TADs change to facilitate lineage-specific gene regulation (Bonev et al., 2017; Rao et al., 2014). TADs substantially constrain enhancer-promoter interactions and multiple studies have shown that the vast majority of enhancer-promoter interactions occur within TAD boundaries (Bonev et al., 2017; Dixon et al., 2012). TAD boundaries are maintained through the binding of CTCF and cohesin as well as specific histone modifications and mark insulated neighborhoods that limit enhancer-gene contacts (see below) (Dixon et al., 2015; Dowen et al., 2014; Rao et al., 2017). TADs have also been shown to localize to particular nuclear compartments, commonly denoted by the letters A and B, where A is the "active" compartment associated with transcriptional activity and accessible chromatin, whereas B is associated with inactivity and repressed chromatin (Dixon et al., 2012; Lieberman-Aiden et al., 2009). Higher-resolution studies have revealed further sub-compartments within these classifications that coincide with replication timing and histone modification patterns (Rao et al., 2014).

Within TADs, poised enhancers can also interact with target genes, suggesting that additional factors or events are needed to activate gene expression (Jin et al., 2013). Hi-C-based studies on enhancer-promoter interaction dynamics suggest that in aggregate, enhancer-promoter interactions in a developmental context follow an "instructive" model whereas new enhancer-mediated gene induction in response to stimuli (such as a specific signaling pathway) relies on preformed contacts (a "permissive" framework) (Bonev et al., 2017; Jin et al., 2013; Phanstiel et al., 2017; Rao et al., 2014). Further support for these results was found in 5C (Chromosome Conformation Capture Carbon Copy) experiments performed on mouse ESCs and neural precursor cells (NPCs) (Phillips-Cremins et al., 2013). Importantly, some interactions identified in ESCs were lost concomitant with changes in enhancer histone marks and target gene

expression upon transition to NPCs. Notably, enhancers that produce transcripts (discussed further below) were more likely to participate in cell type-specific interactions. Further studies to test these models are needed particularly in combination with single-cell approaches to delineate clear interaction patterns.

The role of histone modifications in regulating enhancer activity is also relatively poorly understood, but these marks likely recruit important chromatin remodelers and co-activators or co-repressors. For example, H3K4me1 appears to recruit an ATP-dependent chromatin remodeler complex (BAF, the mammalian homolog of SWI/SNF) to enhancer loci (Local et al., 2017). Interestingly, the catalytic activity of the histone methyltransferases that deposit H3K4me1 is dispensable for recruitment of RNAPII to enhancers and subsequent gene expression, but the proteins themselves are required (Dorighi et al., 2017). On the other hand, both p300 and related complexes that catalyze H3K27ac, the histone modification most strongly correlated with active enhancers, and the mark itself have roles in activating gene expression. H3K27ac appears to directly antagonize the activity of Polycomb Repressive Complex 2 (PRC2) in *Drosophila* (Tie et al., 2009) and chemically inhibition of p300 drastically reduces H3K27ac levels at enhancers and putative target gene expression without substantially changing p300 binding (Raisner et al., 2018). It should also be noted that p300 can acetylate many other proteins besides histones with roles in transcriptional regulation (Dancy and Cole, 2014). Further studies are needed to understand the exact contribution of chromatin modifications as cause or consequence of enhancer activation.

Additional studies have recently connected control of cell-type-specific gene expression through RNA polymerase II (RNAPII) pausing, a rate-limiting event in transcription. In metazoans, RNAPII initiates at transcription start sites and then often pauses after transcribing only a few bases, a process that requires negative elongation factor (NELF) and DRB-sensitivity-inducing factor (DSIF) (Adelman and Lis, 2012). Productive transcriptional elongation occurs by phosphorylation of the DSIF-NELF complex, as well as phosphorylation of RNAPII's C-terminal domain by PTEF-b. Emerging evidence links enhancer activation with RNAPII pause-release

(Chen et al., 2017; Galli et al., 2015; Henriques et al., 2018; Liu et al., 2017b; Schaukowitch et al., 2014). In *Drosophila*, genes with paused RNAPII are enriched for developmental functions in the developing embryo (Zeitlinger et al., 2007), although in mammalian cells, these promoters are enriched for genes that are responsive to stimuli (Min et al., 2011). Notably, the establishment of enhancer-promoter loops early in development, prior to enhancer activation and gene expression, is associated with paused RNAPII in *Drosophila* (Ghavi-Helm et al., 2014). Thus, the regulation of promoter-proximal pausing may be a key aspect of precise enhancer control of gene expression, particularly during development when genes must be activated in a coordinated manner.

Active enhancers produce non-coding transcripts that function in gene regulation

Growing evidence suggests that non-coding transcripts originating from enhancers, called enhancer RNAs (eRNAs), are critical to the function of some enhancers (Arnold et al., 2019; Li et al., 2016). Enhancer transcription is highly correlated with other markers of enhancer activity like H3K27ac and H3K4me1 (Andersson et al., 2014; Zhu et al., 2013) and TF binding (Azofeifa et al., 2018) suggesting that these transcripts might be an indicator of enhancer activity. It should be noted that the exact nature and functional role of eRNAs is a matter of some debate. Many studies characterize eRNAs as short, non-polyadenylated transcripts that are relatively unstable are transcribed bidirectionally (Andersson et al., 2014). However, single-cell transcriptomics have recently shed some light on this phenomenon; a version of CAGE (cap analysis gene expression) in single cells revealed that enhancers that appeared to be bidirectionally transcribed in bulk are actually unidirectionally transcribed from either strand in individual cells (Kouno et al., 2019). There are also eRNAs that more closely resemble long non-coding RNAs (lncRNAs), a broad class of non-coding transcripts with roles in gene regulation (Hsieh et al., 2014; Kowalczyk et al., 2012; Orom et al., 2010) including one characterized as contributing to cardiac hypertrophy (Mirtschink et al., 2019). In most cases, enhancer transcription is thought to regulate gene expression in *cis*; however, evidence also supports a role for eRNAs acting in *trans*, independently of the DNA template they are produced from (Alvarez-Dominguez et al., 2017; Zhao et al., 2016). Given the traditional understanding of enhancers as *cis*-regulatory

elements, these “eRNAs” should perhaps be classified as non-coding RNAs produced from regions that don’t bear the marks of typical promoters, distinct from transcripts produced from enhancers. Together, these results suggest that some “enhancers” function as lncRNA templates, rather than more directly activating transcription via the mechanisms described above.

Enhancer transcription can be stimulated in response to the binding of signal-dependent TFs in order to facilitate enhancer-mediated gene activation (**Figure 3**) (Hsieh et al., 2014; Li et al., 2013). For example, eRNAs may serve as scaffolding for the assembly of TFs, Mediator, and cohesin, and other cofactors to activate target genes (Aguilo et al., 2016; Lai et al., 2013; Pezone et al., 2019; Sigova et al., 2015; Tan et al., 2019; Zhao et al., 2019) and might also contribute to the formation of enhancer-promoter loops (Hah et al., 2013; Kim et al., 2018b; Melo et al., 2013). This process of eRNA-mediated induction of enhancer looping and activation of target genes appears to depend in some cases on the Integrator complex, a protein complex involved in small RNA processing (Baillat et al., 2005; Lai et al., 2015). In some contexts, this complex is required for eRNA termination by RNAPII, a role it also performs at some paused promoters (Barbieri et al., 2018; Elrod et al., 2019). Moreover, eRNAs could play roles in RNAPII promoter-pause release by serving as a decoy for the NELF complex and/or activating the elongation factor PTEF-b (Schaukowitch et al., 2014; Shii et al., 2017; Zhao et al., 2016). This latter role remains speculative since the short RNA produced is not likely sufficient to interact with NELF.

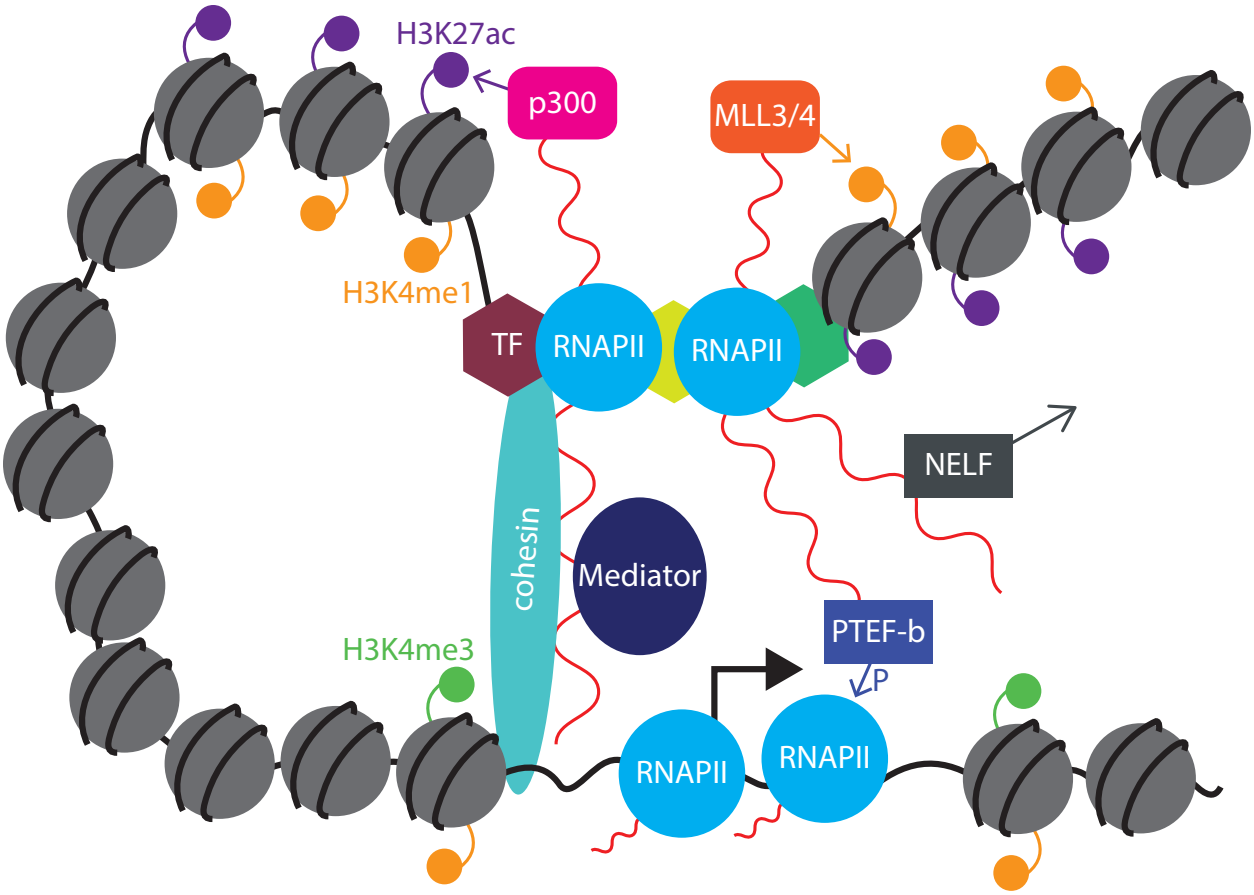


Figure 3: Enhancer RNAs Perform Many Functions Related to Gene Activation
Enhancer RNAs (eRNAs) are involved in many functions related to enhancer-mediated activation of gene expression. eRNAs stimulate the action of chromatin modifiers that deposit enhancer-associated histone marks (top), serve as scaffolding for the assembly of Mediator and cohesin to form enhancer-promoter loops (bottom left), and promote promoter pause-release by sequestering NELF and/or activating PTEF-b.

Reports also suggest that eRNAs modulate the local epigenetic landscape. One study found that blocking eRNA elongation reduced the deposition of H3K4me1 and H3K4me2 at putative enhancer regions during macrophage stimulation and that this effect phenocopied knockdown of the associated methyltransferases (Kaikkonen et al., 2013), although knockdown of other eRNAs had no effect on H3K4me1 levels (Pnueli et al., 2015). Given recent reports discussed above (Dorighi et al., 2017; Local et al., 2017), eRNAs in this context might influence downstream gene expression by recruiting other factors or by modulating the chromatin state of their corresponding enhancers. There are several examples of eRNA knockdown reducing H3K27ac signals at the corresponding enhancer (Bose et al., 2017; Liang et al., 2016; Pnueli et al., 2015). eRNAs have also been found to bind p300 and directly stimulate deposition of

H3K27ac at specific regions (Bose et al., 2017) or recruit chromatin remodelers like CHD1 (Pnueli et al., 2015). The bromodomain-containing protein BRD4 is strongly associated with enhancer activity through binding to acetylated histones (Lovén et al., 2013; Zhang et al., 2012), as it promotes elongation of protein-coding and enhancer transcripts (Kanno et al., 2014). Notably, BRD4 is highly enriched at active enhancers and specifically at super-enhancer clusters that produce eRNAs and have high levels of H3K27ac (Lovén et al., 2013). eRNAs have also been shown to recruit AGO1, a protein more commonly linked with modulating gene expression through small RNA-mediated interference, to help increase the level of long-range interactions within topological domains (TADs) and thus increase downstream gene expression (Shuaib et al., 2019). These data suggest that eRNA production might be part of a positive feedback loop that promotes the interaction of key factors to facilitate downstream gene activation. The fact that eRNAs appear to be involved in many of the phenomena linked to enhancer activity strongly suggests that enhancer transcription is central to enhancer-mediated gene regulation. Future studies will be required to resolve whether enhancer transcription is a cause or effect of activation.

Super-enhancers represent a distinct class of enhancers

Recent work has identified clusters of putative enhancers termed “super-enhancers” (SEs), as large genomic domains consisting of many enhancer elements in proximity to one another (Wang et al., 2019). SEs are recognized by high levels of H3K27ac and the binding of cell-type-specific TFs and/or the Mediator complex (Hnisz et al., 2013; Lovén et al., 2013; Whyte et al., 2013). SEs also tend to reside nearby or overlapping the TSSs of cell-type-specific TFs that bind to the enhancers themselves, forming auto-regulatory positive feedback loops with these key cell-identity regulators (Hnisz et al., 2013; Saint-André et al., 2016). These clusters are distinct from “stretch enhancers”, which are enhancer domains defined primarily by their large size rather than their high total signal of active enhancer hallmarks (Parker et al., 2013). Super-enhancers have been identified in a wide array of mammalian cell types (Jiang et al., 2018; Khan and Zhang, 2015) and are highly specific to cell identity (Wang et al., 2019); even macrophages in different tissue environments show substantial divergence in active super-enhancers

(Gosselin et al., 2014). The idea that super-enhancers represent a distinct paradigm in enhancer-mediated gene regulation remains controversial with some researchers positing that these observations may simply be the strongest examples of active enhancers (Dukler et al., 2016; Halfon, 2019; Hay et al., 2016; Shin et al., 2016). Large enhancer clusters also resemble locus control regions identified decades ago (Levings and Bungert, 2002), and advances in high-throughput sequencing have perhaps simply allowed identification of similar regions genome-wide.

However, recent insights seem to demonstrate that SEs indeed have distinct functions compared to typical enhancers. SEs tend to lie in self-contained "insulated neighborhoods", genomic regions enclosed by CTCF and cohesin, along with the cell-identity genes they regulate more often than typical enhancers (Dowen et al., 2014; Hnisz et al., 2016). This suggests that SEs are organized to robustly activate specific target genes while avoiding inappropriate activation of other genes. Moreover, the high concentration of TFs and transcriptional co-activators at super-enhancer loci appears to create "membraneless organelles" where robust, high levels of transcription can be maintained in "condensates" that promote robust gene expression (Hnisz et al., 2017; Petrovic et al., 2019; Sabari et al., 2018). Condensates appear to require proteins with intrinsically disordered regions (IDRs) which are commonly found in the activation domains of TFs (Nair et al., 2019; Staby et al., 2017) and co-activators like MED1 (a Mediator subunit) and BRD4 (a bromodomain-containing protein known to associate with enhancers) that associate with each other to produce phase-separated regions of high transcriptional activity (Sabari et al., 2018). These results suggest that super-enhancers utilize a distinct regulatory paradigm for controlling cell-type-specific gene expression in a robust manner.

Chromatin accessibility marks regions dynamically occupied by transcription factors

DNA accessibility plays a crucial role in the ability of TFs and the transcriptional machinery to recognize sequence features at promoter and enhancer regions. The vast majority of DNA is encompassed within nucleosomes, octamers of histones wrapped by about 147 base-pairs of

DNA (McGinty and Tan, 2014). Nucleosome dynamics can be influenced by particular post-translational histone modifications (also discussed above), the exchange of canonical histones with histone variants, or by ATP-dependent chromatin remodelers. There is an inverse correlation between nucleosome occupancy and nucleosome turnover at *cis*-regulatory elements with the most accessible regions also having the highest turnover (Klemm et al., 2019).

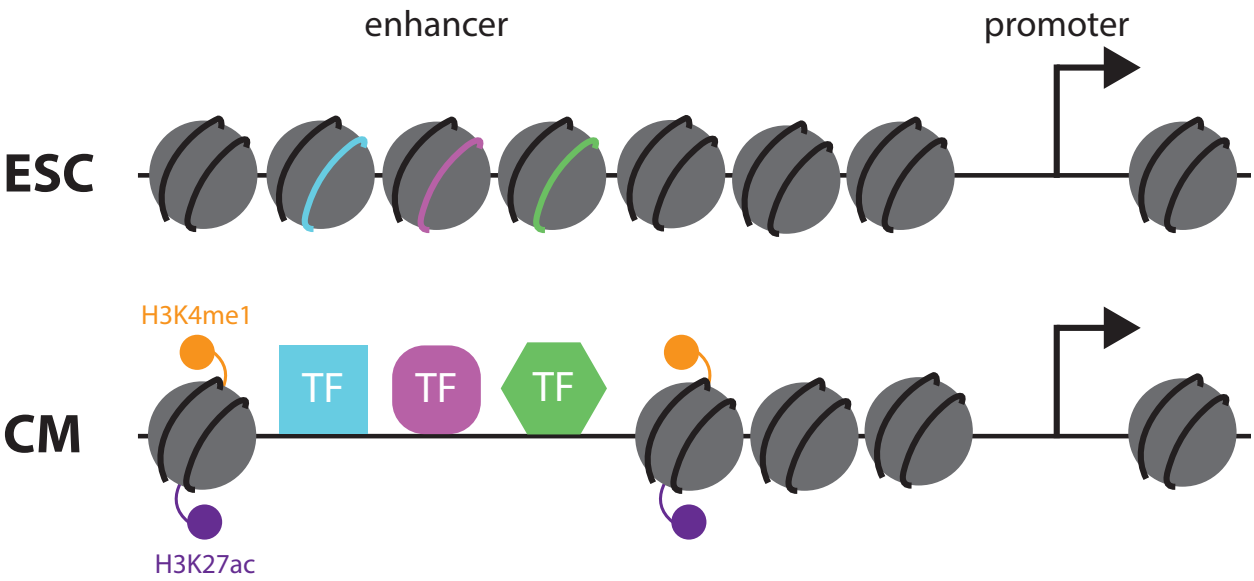


Figure 4: Dynamic Chromatin Accessibility Marks Regions of Lineage-Specific TF Binding
Lineage-specific TFs bind at enhancers (left) which have more dynamic changes in accessibility during development than at promoters (left). Motifs corresponding to these TFs are unavailable for binding due to being packed tightly around nucleosomes (top).

DNA accessibility correlates with TF binding at these regions, and thus measuring changes in open chromatin can reveal the dynamics of TF occupancy at enhancers (**Figure 4**). Although nucleosomes are generally thought to act as a barrier to DNA binding, some TFs prefer to bind nucleosomal DNA. For example, pioneer factors can bind nucleosomal DNA to promote accessible chromatin, sometimes in the absence of ATP-dependent chromatin remodelers (Cirillo et al., 2002; Iwafuchi-Doi, 2018; Mayran and Drouin, 2018). These TFs are discussed in more detail below, but pioneer factors appear necessary to open tissue-specific enhancer regions to allow the binding of other TFs during development or reprogramming.

Altering the composition of nucleosomes by exchanging canonical histones for non-allelic variants is also an important mechanism for regulating chromatin accessibility. Histone variants are distributed to specific loci throughout the genome by ATP-dependent chromatin remodeling complexes and variant- and context-specific chaperones (Talbert and Henikoff, 2016). For example, histone variant H3.3 is deposited by either the chaperones ATRX/DAXX or HIRA (Goldberg et al., 2010). HIRA partners with a single-stranded DNA binding protein involved in DNA replication, RPA, to deposit it at enhancers (Zhang et al., 2017). Studies designed to measure histone dynamics in real time showed that H3.3 turnover is highest at enhancer regions and in super-enhancer clusters (Deaton et al., 2016; Kraushaar et al., 2013). Nucleosomes containing both H3.3 and the histone variant H2A.Z are enriched at *cis*-regulatory elements, including cell-type-specific enhancer regions (Jin et al., 2009). These less-stable nucleosomes can then be evicted by TFs like the retinoic acid receptor to facilitate enhancer activation via binding of other factors. With regards to H2A.Z alone, there is conflicting evidence on its role at enhancers. Contrasting studies performed in the context of estrogen receptor-mediated gene activation suggested that H2A.Z exchange by TIP48 is required for activation of the cyclin D1 gene, but also that at some enhancers H2A.Z is required for enhancer transcription and the recruitment of cohesin (Brunelle et al., 2015; Dalvai et al., 2013). These seemingly disparate results may simply reflect different functional roles of H2A.Z during development (Subramanian et al., 2015). Thus, the deposition of histone variants at specific enhancers can help create a chromatin landscape conducive to gene activation.

Initial studies used DNaseI, an enzyme capable of cleaving DNA in regions of open chromatin, combined with high-throughput sequencing (DNase-seq) to identify regions of accessible DNA on a genome-wide level (Boyle et al., 2008). DNaseI-hypersensitive sites (DHSs) were first identified in *Drosophila* at heat-shock gene promoters, correlating with transcription (Elgin, 1981; Wu et al., 1979b, 1979a). These sites have now been shown to discretely mark TF binding at regulatory elements such as enhancers, silencers, boundary elements and promoters, among others, in a wide array of mammalian cells (Stergachis et al., 2013; Thurman et al., 2012). Among the 49 cell types assayed in one survey, approximately 160,000 DHSs were discovered

per cell type and these DHS signatures could be used to predict the developmental identity of a specific cell (Stergachis et al., 2013). For example, three stages of *in vitro* cardiac differentiation surveyed showed that the broad pattern of DHS sites in ESCs gradually transitioned to a restricted set specific to differentiated CMs, consistent with the idea that DNA accessibility marks cell-type-specific regulatory regions.

Despite the power of DNase-seq in identifying sites of regulatory activity, this approach is limited by experimental constraints such as the requirement of relatively large cell numbers. The Assay for Transposase-Accessible Chromatin (ATAC-seq) uses a modified Tn5 transposase that loads sequence adaptors in regions of accessible chromatin, which has made it possible to survey a much wider array of cell types, particularly those for which collecting large cell numbers is prohibitive (Buenrostro et al., 2013). Comparisons between ATAC-seq and DNase-seq demonstrate a high degree of overlap between accessible chromatin regions identified in the two methods (Buenrostro et al., 2015a; Corces et al., 2017). ATAC-seq has been adapted and combined with expression analysis to learn about gene regulatory networks involved at different stages of cardiac lineage commitment using an hiPSC model (Liu et al., 2017a). This protocol was used to similar effect in CPs in zebrafish to identify enriched TFs involved in cardiogenesis, and these data were extrapolated to identify conserved enhancers in human heart development (Yuan et al., 2018). The ability to measure accessibility at near-base-pair resolution can also be used to identify TF binding events (Bentsen et al., 2020; Gusmao et al., 2014; Pique-Regi et al., 2010; Sherwood et al., 2014), potentially allowing high-throughput assessments of TF binding across the genome in an unbiased manner.

Recent improvements to the ATAC-seq protocol now allow assessment of open chromatin in single cells (Buenrostro et al., 2015b; Cusanovich et al., 2015). This technique, in combination with single-cell epigenomic and transcriptomic techniques, has greatly enhanced our ability to precisely characterize the regulatory complexity of cell identity. For example, a study performed on murine CPs showed that the gene expression of separate populations of progenitors is regulated in overlapping but distinct ways (Jia et al., 2018). These technologies

can also be applied to understand the drivers of disease states. For example, ATAC-seq and gene expression data were combined to characterize the different cell stages of the hematopoietic hierarchy and found that putative enhancer regions delineate lineage progression (Corces et al., 2016). Using this insight, the authors then probed the accessible chromatin landscape of cells from distinct stages of acute myeloid leukemia (AML). From these data, disease states could be assigned to a particular hematopoietic developmental stage, demonstrating how AML co-opts of normal developmental processes. Characteristic TFs of these developmental stages could then be linked to oncogenic states. Together, these studies demonstrate that measuring dynamic chromatin accessibility can uncover the TFs and TF linkages that mediate gene regulation in development and disease.

Transcription factors regulate gene expression

Transcription factors (TFs) are a class of DNA binding proteins that regulate gene expression in all cells, from prokaryotes to eukaryotes (Jacob and Monod, 1961). These proteins often bind DNA directly to sequence-specific motifs, and in eukaryotes they can help recruit the transcriptional machinery including RNAPII to specific regulatory elements including enhancers and promoters, although some TFs act to repress gene expression. In fact, in some cases, the same TF can act as either an activator or a repressor depending on the cell type (Ernst et al., 2016). TFs are typically characterized by one of a diverse set of DNA binding domains. A recent attempt to curate a comprehensive list of every TF in the human genome found only a small fraction with no known DNA binding domain (Lambert et al., 2018). Of this set, about three-quarters had a known binding motif, mostly identified by high-throughput *in vitro* experiments like SELEX (systematic evolution of ligands through exponential enrichment). Such experiments, in which TFs are added to a pool of random DNA sequences that go through multiple rounds of selection, remain the gold standard method for characterizing TF binding motifs (Jolma et al., 2013). Motifs are often described by position weight matrices (PWMs), in which each position has preference scores for each of the four nucleotides (Stormo and Zhao, 2010). The concept of a sequence logo, first conceived to represent consensus sequences identified through genome

alignments, has been repurposed to allow graphical representation of PWMs (Schneider and Stephens, 1990).

PWM-based motifs do not capture all aspects of TF binding, although for most TFs their predictions perform fairly well (Inukai et al., 2017; Slattery et al., 2014; Weirauch et al., 2013). Most motifs span six to ten base-pairs, which means they occur by chance in the genome much more often than actual TF binding sites suggesting that other factors also contribute to binding affinity. The local chromatin environment and DNA methylation also have an impact on TF binding, as discussed above. At the sequence level, models that take into account dependencies between bases, even non-adjacent ones, can improve binding predictions (Zhao et al., 2012). Given the disparity in the structure of TF DNA binding domains, the shape of the DNA molecule might also have substantial impact on sequence recognition. TF binding models that take shape into account along with sequence are generally slightly better at predicting binding than sequence alone (Zhou et al., 2015b). Studies have also shown that protein-protein interactions, either with other TFs or cofactors, can alter TF binding specificity (Garton et al., 2015; Reiter et al., 2017; Slattery et al., 2011). Recent analysis of SNPs in TF genes also showed that even small changes could significantly alter DNA binding specificity and potentially contribute to phenotypic variation between individuals (Barrera et al., 2016; Deplancke et al., 2016). Thus, TF binding is highly complex and involves factors beyond recognition of the primary DNA sequence. Further biophysical and biochemical studies along with computational advances will be needed to fully explain *in vivo* TF binding specificity.

Enhancer sequences are known to harbor many motifs associated with lineage-specific TFs. Multiple models of how TF binding sites are organized at enhancers have been proposed (Long et al., 2016). The “enhanceosome” model, based on early work with the interferon-beta enhancer, suggests that different TFs bind cooperatively at enhancers in precise positions relative to one another to regulate gene expression (Thanos and Maniatis, 1995) (**Figure 5A**). In this paradigm, the tightly defined assembly of a specific set of TFs at a particular enhancer element is required for downstream gene expression in response to viral infection, and thus

their specific binding contexts are tightly constrained. In contrast, the "billboard" model suggests more flexibility, in which enhancers are bound by multiple TFs but not cooperatively (Flores and Ovcharenko, 2018) (**Figure 5B**). An extreme version of this paradigm would suggest that TFs bind at enhancers more or less independently from one another, implying that TF binding is additive in nature, rather than synergistic and cooperative. In actuality, aspects of both of these models may be involved in enhancer regulation. For example, Junion et al. (also discussed above) found that in *Drosophila* heart a TF "collective" of five TFs together regulated the activity of a defined set of cardiac enhancers, but considerable flexibility in motif syntax existed among these sequences (Junion et al., 2012).

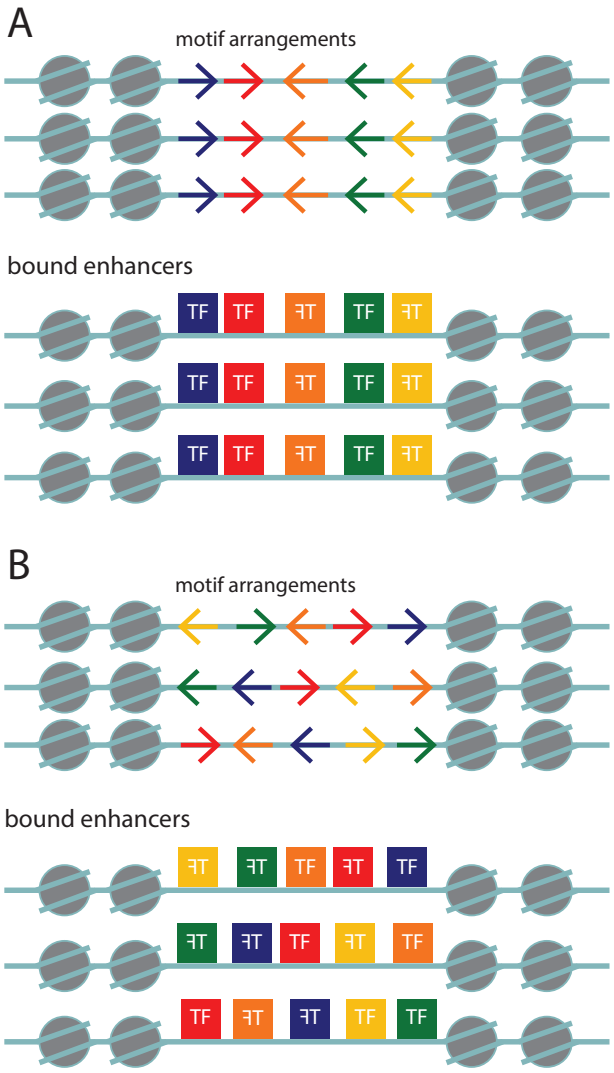


Figure 5: Models of TF Binding at Enhancers

A) Enhanceosome model in which different TFs (indicated by different colors) are fixed in binding position and orientation by the positions of the motifs in the enhancer sequence (arrows).
B) Billboard model, in which a common set of TFs required for activity are not fixed in binding position and orientation relative to one another. Figure partially adapted from Long et al. (2016)

Evidence supports many different mechanisms of TF cooperativity at enhancers. For example, some TFs must cooperate with each other to precisely control spatiotemporal gene expression, like Bicoid in the *Drosophila* embryo (Lebrecht et al., 2005). This cooperativity produces a switch-like behavior, allowing precise patterning as genes are either on or off in a non-linear fashion based on the protein concentration in each nucleus. This behavior is obviously dependent on direct heterotypic or homotypic protein-protein interactions (Jolma et al., 2015; Kerppola and Curran, 1991). Indirect protein-protein interactions can also promote TF cooperativity. For example, a co-factor such as Mediator or p300 that is commonly bound by many classes of TFs and have been shown to facilitate synergistic gene activation (Merika et al., 1998; Spitz and Furlong, 2012). TFs might also assist one another in keeping binding sites available for binding, either by outcompeting nucleosomes, promoting stable chromatin remodeling (pioneer factors), or by modulating DNA shape. Notably, these indirect mechanisms of cooperativity may have less spatial constraints than direct interactions, allowing cooperative binding without requiring tight control of binding site syntax.

As such, recent studies suggest that groupings of relatively low-affinity TF binding sites with loosely-constrained spacing actually allow more cell-type-specific control of gene expression (Farley et al., 2015, 2016; Rickels and Shilatifard, 2018). Low-affinity binding sites also allow control of gene activation by titration of the local TF concentration (Kribelbauer et al., 2019). A comprehensive comparison between active enhancers in a range of tissues in mice and humans showed that the cell-type-specificity of TF binding sites, rather than the specific arrangements of such sites, is most important to enhancer activity, such that loss and gain of particular TF binding sites with different cell-type-specificity could confer different activity profiles (Flores and Ovcharenko, 2018). In fact, the TF binding site composition of promoters is substantially different, featuring more sites overall, a much wider range of TFs, and more overlap between them, leading to far less cell-type-specificity (Mattioli et al., 2019). These observations suggest

the overall composition and density of binding sites at enhancers is constrained in order to ensure context-specific activity.

TF binding at enhancers can have various consequences on the local genomic environment to promote or inhibit gene expression. For example, TFs can influence chromatin accessibility, directly or indirectly. As discussed above, pioneer factors can displace nucleosomes to facilitate the recruitment of other TFs or co-factors that chemically modify histones and/or DNA (Benveniste et al., 2014). For example, DNA methylation, catalyzed by a class of DNA methyltransferases that add a methyl group to cytosines in a CpG dinucleotide context in mammals (Smith and Meissner, 2013), can positively or negatively affect the ability of TFs to recognize and bind their motifs (Maurano et al., 2015; Rishi et al., 2010; Wang et al., 2012). TFs can affect DNA methylation states and thus the ability of other factors to bind by blocking the activity of DNA methyltransferases (Brandeis et al., 1994; Macleod et al., 1994), through the direct recruitment of these same methyltransferases (Quenneville et al., 2011), and/or recruitment of demethylases (Fujiki et al., 2013). TFs can also recruit chromatin modifiers to silence genes (Arnold et al., 2013b; Ecco et al., 2016) or activate them (Mansour et al., 2014), as discused above. On a broader scale, TFs can also modulate genome topology. Beyond the well-known role of CTCF in establishing and maintaining TAD boundaries as well as other genome loops (Ong and Corces, 2014), studies in ESCs have demonstrated that key pluripotency TFs help spatially organize the genome to facilitate the expression of pluripotency genes (Stevens et al., 2017; Wit et al., 2013). Recent work also indicated a role for the TF YY1 in directly linking enhancers and promoters (Weintraub et al., 2017). Collectively, these data demonstrate that TFs use a variety of mechanisms to facilitate proper gene regulation.

Many TFs activate transcription as a direct response to signaling inputs, resulting in some cases in translocation of a factor from the cytoplasm to the nucleus where it can bind *cis*-regulatory elements. Prime examples include of beta-catenin which can bind DNA elements in response to Wnt signaling (Cong et al., 2003), translocation of the estrogen receptor to specific binding sites after ligand binding (Björnström and Sjöberg, 2005), and nuclear localization of the co-

activators YAP and TAZ (discussed below) in response to Hippo signaling. Curiously, in most cases these TF binding events are unable to activate gene expression on their own; this "activator insufficiency" adds further complexity to the picture of developmental gene regulation (Barolo and Posakony, 2002). This scenario allows a particular signaling pathway to activate lineage-specific gene expression in an array of different contexts, partnering with different co-factors and TFs that are stimulated by other signaling pathways, while preventing ectopic activation of genes related to different lineages. These complex gene regulatory networks have to be considered when attempting to study mechanisms of TF action during development – the perturbation of a signaling pathway or its downstream effectors may have fairly complex and non-linear impacts on the genes that they regulate. Multiplex and single-cell technologies along with the associated computational and bioinformatic advancements are necessary to develop a full understanding of the networks of TF action during development (Wilkinson et al., 2017).

Pioneer factors and master regulators are transcription factors that underpin cell identity

As discussed above, pioneer factors are TFs that establish accessible chromatin at *cis*-regulatory elements. Bioinformatic analyses have revealed that motifs corresponding to pioneer factors are overrepresented in enhancer regions (Vandel et al., 2019). Notably, many TFs known to be critical to maintaining stem cell pluripotency have pioneer activity, binding to closed chromatin to help open enhancer regions that are critical to establishing and maintaining the pluripotent stem cell state (Adachi et al., 2018; Friman et al., 2019; Koche et al., 2011; Soufi et al., 2012). These pluripotency-promoting TFs may also help to preserve relative accessibility at enhancers that could be active in downstream developmental lineages (Kim et al., 2018a), suggesting a broader regulatory paradigm in which TFs with pioneer activity act at different stages of development to potentiate enhancer activation (Cernilogar et al., 2019). These data also suggest that pioneer factors play a critical role in development to control cell-type-specific gene expression.

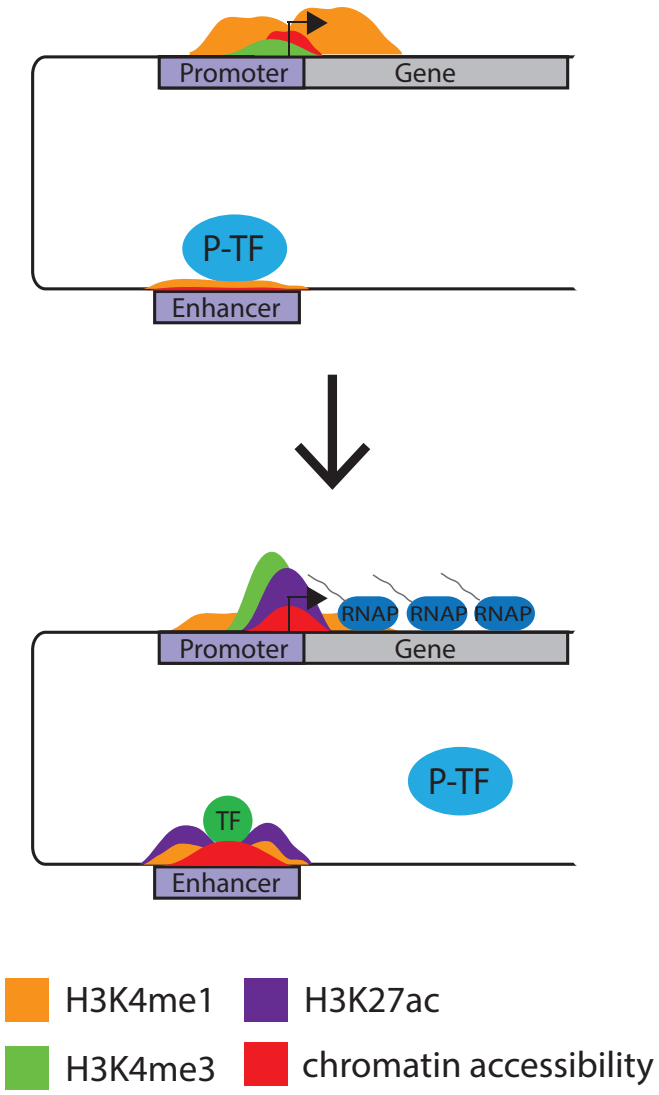


Figure 6: Pioneer Factors Potentiate Enhancer Activation
Cell type-specific transcriptional activation is largely mediated by the binding of TFs to enhancer elements. A pioneer TF (P-TF) binds and initiates chromatin opening to prime the area for future activation (top). Open chromatin allows for the subsequent exchange of the P-TF for other master TFs (bottom). TFs concomitantly recruit chromatin modifiers and other factors that lead to enhancer activation and subsequent gene expression.

Detailed studies have revealed a few potential mechanisms as well as common structural features used by pioneer factors to establish and maintain chromatin accessibility at enhancers (Garcia et al., 2019; Mivelaz et al., 2019). For example, the DNA binding domain of FoxA family members resembles the nucleosome-binding domain of linker histones, which may allow these factors to bind directly to nucleosomes (Clark et al., 1993; Iwafuchi-Doi et al., 2016). Other pioneer factors appear to bind partial binding sites that are revealed during the natural "breathing" that occurs with nucleosome-bound DNA (Huertas et al., 2020; Soufi et al., 2015).

Many pioneer TFs have been shown to directly interact with ATP-dependent chromatin remodelers to subsequently modulate the chromatin environment (King and Klose, 2017; Takaku et al., 2016; Vierbuchen et al., 2017; Wang et al., 2014). Pioneer factors also recruit chromatin modifiers that deposit enhancer-associated histone marks including p300 (Fuglerud et al., 2018) and the MLL complex (Jozwik et al., 2016; McKinnell et al., 2007) as well as induce the loss of DNA methylation (Sardina et al., 2018; Serandour et al., 2011; Vanzan et al., 2020; Zhang et al., 2016). Collectively, these actions can facilitate the binding of non-pioneer TFs to enhancers and thus act as catalysts for downstream gene expression.

The ability of pioneer factors to bind inaccessible chromatin across the genome allows lineage-specific TFs that may lack this ability to cooperate with pioneers to stabilize their binding at specific *cis*-regulatory regions, and thus facilitates further opening to the binding of other factors (Donaghey et al., 2018; Mayran et al., 2019). A recent study in liver development suggested that even transient binding events contribute to the gradual opening of regulatory regions, which primes them for full activation when signaled to do so (Karagianni et al., 2020). TFs that are necessary for the activation of a lineage-specific gene expression program have also been classified as “master regulators”. This term was first coined to refer to a gene that stands at the top of a regulatory hierarchy (Ohno, 1979), and over the years has become associated with TFs that are required for the establishment of specific lineages. Because of this, the use of this term has exploded in recent years, as it turns out that the expression of most TFs is restricted to a small subset of lineages (Chan and Kyba, 2013; Lambert et al., 2018). However, since many of these lineages are downstream from others, it may be hard to argue that the TFs important to the establishment of these identities are truly at the top of a regulatory hierarchy. Recent advances in cellular reprogramming have been instrumental in defining true “master regulators”: TFs that are necessary and sufficient to catalyze a permanent change in cell identity. The discovery that MyoD can induce muscle cell fate in fibroblasts and other non-muscle cells and later work showing that expression of the OSKM TFs (*OCT4*, *SOX2*, *KLF4*, and *MYC*) in fibroblasts is sufficient to induce a reversion to the pluripotent stem cell state indicate that these factors function at the top of TF regulatory hierarchies (Takahashi and Yamanaka,

2006; Tapscott et al., 1988). Notably, whereas pluripotency TFs do appear to have pioneer activity on their own as described above, MyoD needs to partner with another TF that is already bound to target sites (Berkes et al., 2004). These insights suggest that pioneer and master regulatory TFs represent overlapping but distinct roles in gene regulation. In some lineages, a master regulator or a set of master regulators is required to establish and maintain cell identity, whereas pioneer factors may also be required to establish a developmental lineage by marking enhancers for activation, but they may not necessarily be linchpins that drive changes in cell identity on their own.

Despite studies on pioneer factors in developmental contexts, identification of pioneer factors in the cardiac lineage have remained elusive. As discussed above, TFs such as GATA4 and MEF2C catalyze cardiac reprogramming (Ieda et al., 2010), but an ability of MEF2C to bind nucleosome-bound DNA has not been demonstrated. GATA4 appears to have pioneer-like activity in other systems (Bossard and Zaret, 1998; Cirillo et al., 2002), and it can induce progression from mesoderm towards the cardiac lineage when co-expressed with a chromatin remodeler (Takeuchi and Bruneau, 2009). Recent studies have suggested that the TF ISL1, a marker of the second heart field (SHF), acts a pioneer factor (Gao et al., 2019). Although both pioneer factors and master regulators clearly play important roles in establishing and maintaining cell fate, further studies that identify pioneer-like factors in the cardiac lineage are key to enhancing our understanding of the control of gene expression during heart development and disease.

Transcriptional regulation of heart development

The heart is the first functional organ during mammalian embryogenesis. After gastrulation, the cells that become the heart are derived largely from the lateral plate mesoderm (Buckingham et al., 2005). These mesodermal cells are specified by inductive signals from the endoderm (Arai et al., 1997). Lineage progression leads to formation of cardiac mesoderm, marked by the expression of the master regulatory TF Mesp1 (Bondue et al., 2008; Saga et al., 1999) and then cells further specialize into cardiac progenitors that differentiate into a diverse array of cardiac

cell types (**Figure 7**). This entire process relies on the transmission of complex extracellular signaling cues into specific transcriptional outcomes whereby TFs act as the receiver and sender of these signals (Noseda et al., 2011).

Almost all cells comprising the heart are derived from two pools of cardiac progenitors known as heart fields (Meilhac and Buckingham, 2018; Meilhac et al., 2004; Später et al., 2014). The cells in the first heart field (FHF) progress to a cardiomyocyte cell fate earlier than those in the second (SHF). At this stage, these pools of progenitors can be distinguished by the expression of key markers – the FHF progenitors are distinguished by *Hcn4* expression (Liang et al., 2013; Später et al., 2013), whereas the SHF progenitors are marked by expression of *Isl1* (Bu et al., 2009; Moretti et al., 2006). These cells undergo a series of migrations and morphological changes to form an early heart tube, after which looping occurs, followed by growth and morphological changes that eventually form the multi-chambered, fully developed heart (Buckingham et al., 2005). Lineage tracing experiments have demonstrated that FHF progenitors differentiate primarily into CMs of the left ventricle and the atria (Liang et al., 2013; Später et al., 2013), whereas SHF progenitors primarily contribute to the right ventricle and the atria (Bu et al., 2009; Moretti et al., 2006; Später et al., 2013). Both of these progenitor pools also contribute to CMs that form the cardiac conduction system (Mikawa and Hurtado, 2007; Mohan et al., 2018). Canonical Wnt signaling promotes formation of cardiac mesoderm in part through induction of SOX17 (Liu et al., 2007), after which Wnt signaling represses cardiac cell fate (Naito et al., 2006; Ueno et al., 2007). After this key decision point, non-canonical WNT, BMP, and FGF signaling promotes differentiation of FHF progenitors, whereas canonical WNT, sonic hedgehog, and FGF signaling promotes proliferation of SHF progenitors, after which non-canonical WNT and BMP signaling promotes their further differentiation (Galdos et al., 2017). Thus, proper induction of developmental gene expression programs requires the precise cross-talk between signaling pathways and specific TFs during the induction of cardiac cell types that form the three-dimensional structure of the heart.

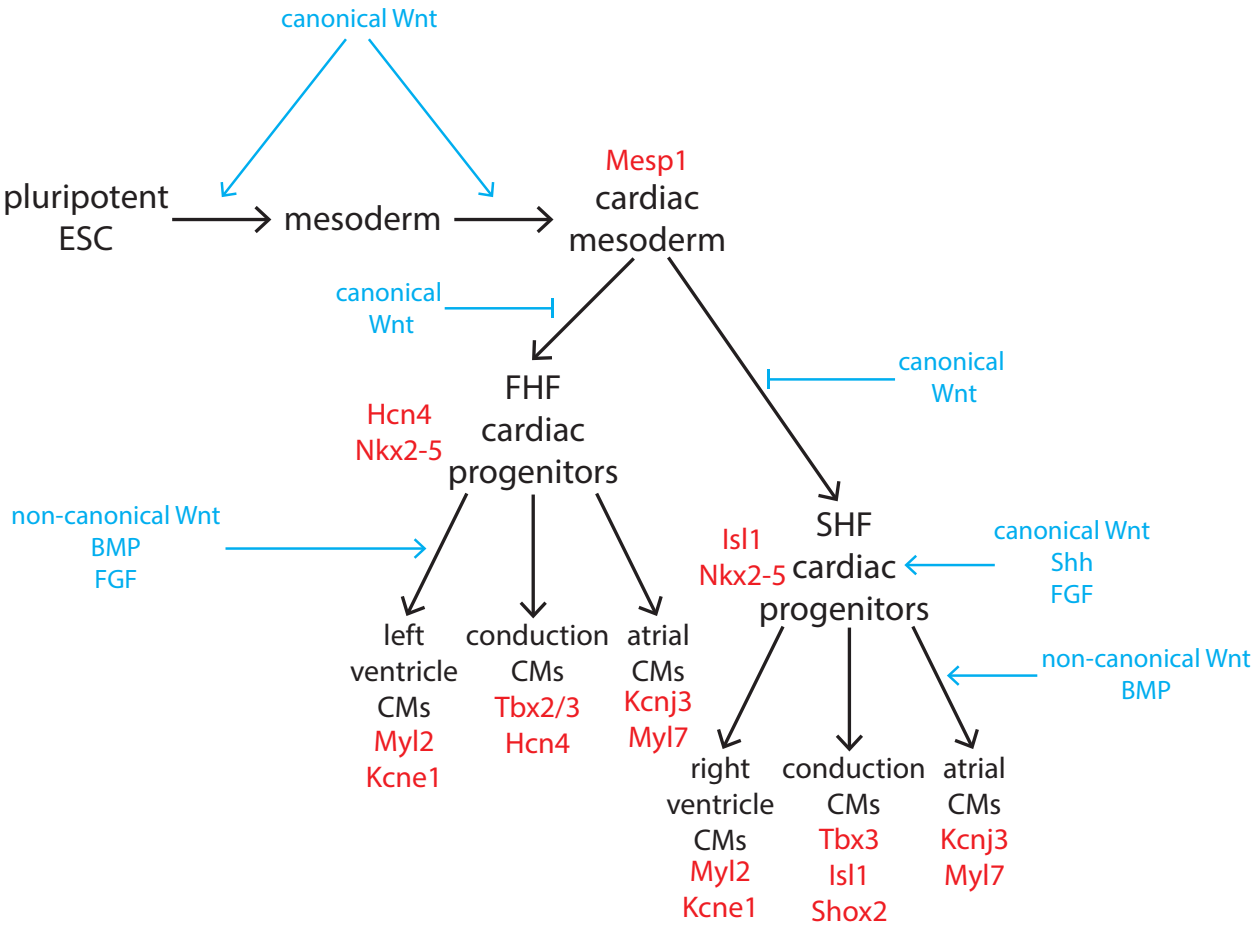


Figure 7: Signaling Inputs Drive the Progression of Cardiac Muscle Lineage Commitment
Depiction of lineage commitment of the myocardium. Pluripotent ESCs progress through mesoderm on to cardiac mesoderm, and then to two pools (heart fields) of cardiac progenitors, which contribute to the specified populations of cardiomyocytes in the adult heart (as well as cells in other heart tissues not shown). Signaling pathways that drive progression or expansion of particular populations shown in light blue, and marker genes of each population shown in red.

CMs undergo a maturation process that continues from mid-fetal development to well past birth. Maturation involves substantial morphological, metabolic, and electrophysiological changes in response to the increased load required for CMs to pump blood throughout the body (Karbassi et al., 2020). Although stem-cell-based approaches provide investigators with an *in vitro* model for studying cardiac differentiation, CMs differentiated from pluripotent stem cells in culture do not appear to reach full maturation even after extended culture (Uosaki et al., 2015). Thus, continued studies into the mechanisms of cardiac maturation will be critical to improve methods for *in vitro* cell culture.

Other cell types perform important roles in the development and function of the myocardium. These tissues include the endocardium, a layer of endothelial cells that lines the inside of the heart muscle; the epicardium, a thin outer layer of epithelial cells that encompasses the myocardium, and the pericardium, a sac that surrounds the whole heart. The endocardium and the myocardium arise from a common pool of progenitors originating from pre-cardiac mesoderm (Devine et al., 2014; Misfeldt et al., 2009; Tam et al., 1997), although it appears that FHF progenitors have mutually exclusive potential fates (endocardium *or* myocardium), whereas SHF progenitors retain multipotency (Lescroart et al., 2014). The endocardial cells contribute to many key tissues in the adult heart (Nakano et al., 2016; Zhang et al., 2018). During heart morphogenesis, some endocardial cells, responding to signals from the myocardium (Zhang et al., 2018), undergo endothelial-to-mesenchymal transition (EndoMT) and migrate through the cardiac jelly (a layer of extracellular matrix between the endocardium and the myocardium) to form endocardial cushions which will eventually fuse and reshape themselves into the heart valves (Lange et al., 2004). Once the developing heart grows to a size that requires vasculature to supply oxygen and nutrients, endocardial cells from the sinus venosus begin to form coronary vasculature in an initial wave of migration into the myocardium, although there are later waves to form vessels in the ventricular septum and the inner myocardial walls that are derived from the ventricular endocardium (Zhang et al., 2018). There are also pericytes, mesenchymal cells that line microvascular blood vessels, and smooth muscle cells, which line larger arteries and veins and can be derived from the epicardium and the endocardium (via EndoMT) (Chen et al., 2016). In addition, cells from the epicardium undergo epithelial-to-mesenchymal transition (EMT) and give rise to most of the fibroblasts in the heart (Ivey and Tallquist, 2016; Wessels et al., 2012). Epicardially-derived cells also contribute to smooth muscle and mesenchymal tissue in heart valves (Lie-Venema et al., 2007; Wessels et al., 2012). These tissues, in addition to forming vital support structures, supply important paracrine signaling cues that promote proper myocardial development and proliferation (Buikema et al., 2013; Kerkela et al., 2008; Li et al., 2011; Später et al., 2014). Of note, recent single-cell transcriptomic studies have confirmed many of these insights into the lineages that form the tissues of the heart (particularly the myocardium), while also uncovering

new linkages between specific tissues and transcriptional regulators (Lescroart et al., 2018; Soysa et al., 2019).

The complexity of this process necessitates precise control of gene expression in time and space as faulty regulation can have deleterious developmental consequences. Congenital heart disease (CHD) is the most common birth defect, manifesting in almost 3% of births worldwide (Gelb and Chung, 2014). Despite this high incidence, the mechanisms leading to CHD are rather poorly understood, although there is a significant link to mutations in cardiac TFs (Bruneau, 2013; Jin et al., 2017; Prendiville et al., 2014). Studies suggest that *de novo* mutations also play a significant role (Homsy et al., 2015; Øyen et al., 2009; Sifrim et al., 2016; Zaidi et al., 2013). Notably, these *de novo* mutations appear enriched for genes encoding chromatin regulators as well as TFs, underscoring the importance of gene regulation to proper heart development. In addition, CHDs can often occur in the context of broader genetic syndromes like Down syndrome suggesting that additional complex mechanisms of genetic regulation underlie a large proportion of cases of CHD (Richards and Garg, 2010; Sifrim et al., 2016). Thus, a clear understanding of the molecular mediators driving cardiac-specific gene expression including cell-type-specific TFs will be required to improve diagnostic and therapeutic strategies.

Genetic studies in mouse and zebrafish have provided important information regarding the key TFs that direct heart development. For example, NKX2-5 and TBX5 both have roles controlling cardiac gene expression through physical interactions at enhancers in CMs (Hiroi et al., 2001; Lints et al., 1993; Lyons et al., 1995). Single- and double-knockouts of these TFs showed that they can function independently to regulate some genes and synergistically to regulate the expression of others, and that these patterns can differ based on developmental stage (CP or CM) (Luna-Zurita et al., 2016). This study revealed that at some binding sites, these TFs are dependent on binding of the other factor, mutually dependent, independent, or require the binding of other factors altogether. These data suggest a great deal of complexity in the regulation of cardiac gene expression.

We now also know that many of the essential TFs involved in heart development control key cardiac-specific genes through their binding at enhancers (Akerberg et al., 2019; Luna-Zurita et al., 2016). In fact, there is strong evidence from human genetic studies suggesting that enhancers control the expression of key genes during heart development and that subtle changes in enhancer sequences can disrupt these expression patterns. For example, Holt-Oram syndrome, a genetic disease characterized by cardiac septation defects and skeletal malformations (Holt and Oram, 1960), has been linked to mutations in the gene encoding TBX5 (Basson et al., 1997). Remarkably, a single base-pair mutation in a cardiac-specific enhancer abrogated *TBX5* expression in a patient with Holt-Oram syndrome (Smemo et al., 2012). More broadly, genome-wide association studies (GWASs) demonstrate that many cardiac trait-linked single-nucleotide polymorphisms (SNPs) reside in non-coding regions that overlap enhancers (Holm et al., 2010; Pfeufer et al., 2010; Postma et al., 2016; Tayal et al., 2017; Wang et al., 2016). For example, GWAS SNPs associated with QT interval length, a measure of myocardial repolarization, are highly enriched in cardiac-specific enhancers active in the left ventricle (Wang et al., 2016). Thus, heart development clearly requires tight spatiotemporal coordination of gene expression in response to signals, and enhancers are key to that process through translating TF binding into transcriptional outputs.

Biochemical and mechanical cues target TEAD TFs to drive gene expression

The Hippo signaling cascade was first characterized in *Drosophila melanogaster* (Wu et al., 2003). The kinase Hpo (MST1/2 in mammals) phosphorylates and activates the kinase Wts (LATS1/2 in mammals) with the help of the scaffolding protein Sav (SAV1 in mammals) (Avruch et al., 2012) (**Figure 5**). This study and others determined that the Hpo kinase cascade acted to negatively regulate proliferation through attenuating the expression of cell cycle and apoptosis-related genes (Tapon et al., 2002; Wu et al., 2003). Further work identified Yorkie (Yki), the *Drosophila* homolog of the mammalian YAP and its paralog TAZ (Huang et al., 2005) as the terminal target of this kinase cascade. When Hippo signaling is active, phosphorylation of Yki prevents its nuclear localization partially through the action of 14-3-3 proteins (in mammals, 14-3-3σ) and thus, prevents it from activating the expression of proliferation genes (Dong et al.,

2007; Ren et al., 2009; Zhao et al., 2007). When Hippo signaling is off, Yki can translocate to the nucleus to interact with the TF Scalloped (Sd) (TEAD1-4 in mammals, discussed below) to promote proliferation (Goulev et al., 2008). In mammals, inhibition of Hippo signaling leads to nuclear import of YAP and binding to TEAD TFs to regulate cell growth (Ota and Sasaki, 2008). Hippo signaling has been linked to control of organ size and various other developmental roles, such as in the liver and the intestine, and its misregulation is associated with a variety of cancers (Fu et al., 2017; Harvey et al., 2013; Mo et al., 2014; Pan, 2007; Yu et al., 2015). More recently, Hippo signaling has been linked to heart development and maintenance of homeostasis in adult hearts (Zhou et al., 2015a).

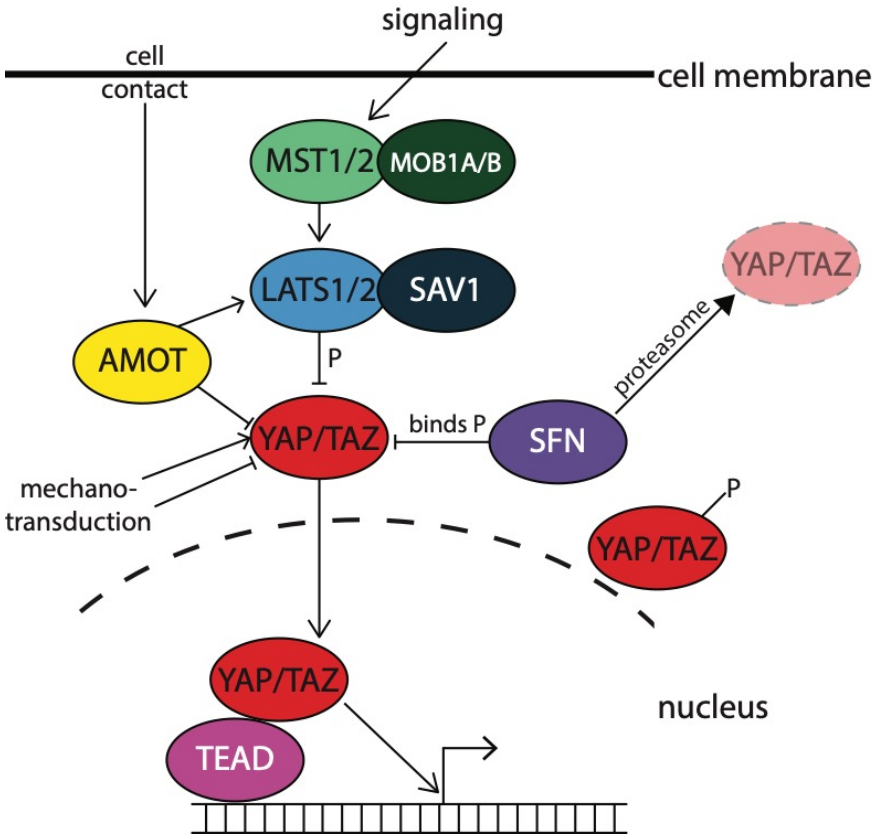


Figure 8: Hippo Signaling Translates Diverse Cues into Activation of Gene Programs
MST1/2, with the assistance of the scaffolding protein MOB1A/B, phosphorylates LATS1/2. LATS1/2, with the assistance of the scaffolding protein SAV1, phosphorylates the transcriptional co-activator YAP and/or TAZ. These phosphorylated proteins are then targeted for degradation by SFN (14-3-3σ), although nuclear exclusion is also independent of SFN-mediated degradation. Angiomotin family proteins (AMOT) also can activate LATS1/2 phosphorylation of YAP/TAZ, as well as sequester YAP it at the membrane in response to cell-cell contact. Mechanotransduction can also influence YAP/TAZ subcellular localization in response to specific cues. Unphosphorylated, YAP and TAZ can enter the nucleus and activate TEAD target genes.

Specific ablation of various components of the Hippo signaling cascade including *Salv*, *Mst1/2*, and *Lats2* results in perinatal lethality in mice with enlarged hearts and a higher number of CMs compared to wild-type mice (Heallen et al., 2011). Notably, this phenotype was partially rescued by reducing the levels of beta-catenin (the downstream effector of canonical WNT signaling) since YAP and TEAD appear to form a complex with beta-catenin at the promoters of some proliferation-associated genes. YAP mutants that prevent its phosphorylation in CMs also led to excess CM proliferation and enlarged hearts, whereas YAP knockouts caused the opposite phenotype (Gise et al., 2012; Xin et al., 2011). Thus, precise regulation of Hippo signaling is critical to proper heart development, converging with TEAD TFs to drive changes in gene expression.

Consistent with this idea, a study in zebrafish showed expression of dominant-negative forms of Yap and Taz/Wwtr1 unable to interact with Tead factors blocked CPs from properly migrating to the midline, although these cells maintained the ability to differentiate into CMs (Miesfeld and Link, 2014). More recent work in zebrafish showed that *wwtr1* knockouts also have defects in cardiac trabeculation as a result of disordered cortical actin structure and irregular cell-cell junctions (Lai et al., 2018), a defect that mirrors the mouse phenotype of *Tead1* loss (Chen et al., 1994). However, *Taz* knockout mice have broad defects outside the cardiovascular system (Hossain et al., 2007; Makita et al., 2008). Outside of the myocardium, *Lats1* and *Lats2* deletion in the epicardium leads to defects in coronary vasculature and differentiation into cardiac fibroblasts (Xiao et al., 2018). These examples suggest a link between cardiac morphogenesis and gene expression regulated by Hippo signaling through TEAD factors. Nevertheless, TEAD TFs have not been well studied in heart. YAP and TAZ have also been linked to CHD due to their ability to act as transcriptional co-activators with the cardiac TF TBX5 (Murakami et al., 2005). YAP and TAZ can also act as co-activators with a wide array of other TFs, some of which also have roles in cardiac development (Wang et al., 2018).

The activity of YAP can also be regulated in ways independent of the canonical Hippo cascade but connected to mechanical stimuli from the cellular environment. For example, the

angiomotin (AMOT) family of proteins can regulate cell growth and motility by activation of LATS1/2 independently of MST1/2 through co-localization at tight junctions (Paramasivam et al., 2011). Given the importance of cell-cell contacts for cardiac contraction, it is perhaps unsurprising that this non-canonical Hippo pathway regulates CMs. In cardiac progenitors, mechanical sensing of stiffness of the extracellular matrix appears to drive cell fate decisions between CMs and endothelial cells, and also controls the nuclear entry of YAP/TAZ. Silencing YAP abrogates this lineage decision suggesting that mechanosensing-driven nuclear entry of YAP/TAZ is key to cardiac lineage commitment (Mosqueira et al., 2014). The proto-cadherin FAT4 appears to be required in CMs to sequester the angiomotin protein AMOTL1, which associates with YAP1 at cell-cell junctions. In the absence of FAT4, AMOTL1 and YAP1 enter the nucleus and promote proliferation (Ragni et al., 2017). Merlin, a member of the protein 4.1 superfamily that links the actin cytoskeleton to the plasma membrane, is activated in arrhythmogenic cardiomyopathy, causing activation of the Hippo kinase cascade and reduced expression for TEAD targets (Chen et al., 2014; Zhang et al., 2010). Recent work in non-cardiac cell models adds further support for the idea that the subcellular localization of YAP and its association with TEAD TFs can be influenced by mechanotransduction (Chang et al., 2018; Ege et al., 2018; Elosegui-Artola et al., 2017). Given the prominent role that actomyosin structures play in the force-generating function of cardiac muscle cells, such insights suggest that TEAD TF activity might be specifically controlled by interactions between CMs and the extracellular matrix or in response to stress. Thus, mechanical cues are important for proper development and functioning of the heart muscle and that this information may be translated into transcriptional regulation at least in part by influencing localization of YAP and its ability to bind TEAD to activate gene programs.

Hippo signaling also has important roles in postnatal homeostasis of the heart. The adult human heart has a relatively low capacity to generate new CMs as a result of the limited postnatal proliferative capacity of CMs (Senyo et al., 2013). Postnatal cardiomyocyte-specific deletion of Hippo cascade components (*Salv* and *Lats1/2*) restored the proliferative capacity of cardiomyocytes in murine hearts (Heallen et al., 2013). In fact, a micro-RNA (miRNA)-based for

stimulation of proliferation in iPSC-derived cardiomyocytes recovered many miRNAs that repress Hippo cascade components (Diez-Cuñado et al., 2018). Conversely, *Mst1* has been shown to promote cardiomyocyte apoptosis when activated by oxidative stress (Del Re et al., 2014). Remarkably, CM-specific deletion of *Salv* after myocardial infarction-induced ischemic heart failure allowed renewal of cardiomyocytes and restoration of heart function (Leach et al., 2017). However, long-term Hippo inhibition through CM-specific *Salv* knockdown led to eventual diminution of cardiac function under transverse aortic constriction, despite the appearance of some hallmarks of regenerative activity (decreased cell death and cell cycle re-entry) (Ikeda et al., 2019). Taken together, these results suggest that blocking Hippo signaling might allow adult cardiomyocytes to revert to a less mature state to allow cell cycle re-entry, but that its re-activation is also likely necessary for maturation into a fully functional heart.

TEAD TFs have roles in cell growth and cardiomyocyte development

Extracellular signaling drives developmental gene expression by targeting TFs especially at enhancers (Noseda et al., 2011). As discussed above, the Hippo kinase cascade is thought to modulate the subcellular localization of the transcriptional co-activators YAP1 and its homologs as well as its paralog TAZ to bind to TEAD TFs in the nucleus (Goulev et al., 2008; Ota and Sasaki, 2008). The TEAD family consists of four members, TEAD1-4. TEAD1 (also known as TEF-1 for transcriptional enhancer factor 1), the archetypal family member, was identified as a cell-type-specific factor that bound to the SV40 viral enhancer (Davidson et al., 1988; Xiao et al., 1987). TEAD1 was subsequently characterized as a TF linked to enhancer-mediated gene regulation, yet it lacked proper activation domain, explaining its need for co-activators like YAP/TAZ (Xiao et al., 1991). TEAD1 binds to a canonical motif through a highly conserved DNA binding domain known as the TEA domain (Bürglin, 1991). The identification of this domain has allowed the identification of other TEAD family members in mammals.

gene	KO phenotype	reference(s)
Tead1	embryonic lethality (E11-12) due to heart and brain defects	Chen et al. (1994), Sawada et al. (2008)
Tead2	some prenatal lethality due to neural tube closure defects	Kaneko et al. (2007)
Tead3	viable with skeletal and metabolic defects	Dickinson et al. (2016)
Tead4	pre-implantation lethality due to failure to form blastocoel	Nishioka et al. (2008)

Table 2: Knockout Mouse Phenotypes of TEAD Family Members

TEAD TFs have been linked to control of gene expression in a variety of contexts from development (Dickinson et al., 2016; Kaneko et al., 2007; Milewski et al., 2004; Nishioka et al., 2008; Sawada et al., 2008; Yagi et al., 2007) to cancer (Pobbati and Hong, 2014). As shown in **Table 2**, TEAD factors play crucial roles in mammalian embryonic development, with TEAD1 most strongly implicated in heart development. Other TEAD family members also appear to have specific roles in controlling cardiac-specific gene expression. TEAD3 (also known as DTEF-1) was shown to regulate cTnT through binding at its promoter in chick embryonic cardiac development as well as to interact with MEF2 TFs (Azakie et al., 2005). TEAD3, TEAD4, and TEAD1 can also interact with the cardiac muscle TF MEF2C to differentially modulate the expression of MEF2-dependent genes (Maeda et al., 2002a). These studies suggest that TEAD factors play an underappreciated role in regulating heart development through the control of cardiac-specific genes, and that these family members may perform distinct roles depending on cellular context.

TEAD TFs are also notably connected to extracellular signaling in CMs independently of Hippo signaling. *Tead1*, along with C/EBPβ, is downregulated by p38 MAPK signaling in cardiomyocytes preventing activation of certain target genes (Ambrosino et al., 2006; Ben-Levy et al., 1998). Notably, α_1-adrenergic signaling, which stimulates cardiac hypertrophy and can

reactivate the fetal cardiomyocyte gene expression program (Simpson, 1983; Waspe et al., 1990) activates *Tead3* and *Tead4*, but not *Tead1* (Maeda et al., 2002b; Stewart et al., 1998). In addition, chelation of nuclear calcium ions appears to increase Tead transcriptional activity (Thompson et al., 2003),although these experiments were not performed in CMs, calcium signaling has a major role in controlling gene expression and function in cardiac muscle.

Of the TEAD family members, TEAD1 is most strongly linked to cardiac-specific gene regulation during development. *Tead1*-knockout mice die during embryonic development between E11 and E12 showing cardiac underdevelopment (Chen et al., 1994), however, the exact phenotype and its causes are still unclear. In other KO studies of *Tead1* and *Tead2*, mice showed broad defects in proliferation and a failure to fuse heart tubes (Sawada et al., 2008). However, these studies were limited by conventional knockout strategies rather than conditional loss of TEAD1 during different stages of heart development. More recent work has also shed some light on TEAD1's potential role in heart development. He et al. identified the TEAD motif as highly enriched in cardiac enhancers in HL-1 cells and confirmed TEAD1 binding by ChIP-qPCR to selected enhancers (He et al., 2011). In follow-up work, TEAD1 was shown to bind to an array of cardiac-specific enhancers along with a set of other cardiac TFs. The authors indirectly leveraged gene expression data in *Tead1*-KO CMs to suggest that this TF plays a role in cardiac muscle contractility and structure (Akerberg et al., 2019). Thus, this study was unable to examine the direct consequences of TEAD1 binding on enhancer activity during cardiac lineage commitment. This thesis identified a direct role for TEAD1 in the dynamic establishment of enhancer states during cardiac lineage commitment, as I will describe in Chapter 2. Together, prior studies combined with our current work has revealed novel insights into enhancer regulation during cardiac lineage commitment and demonstrates an unbiased approach to uncover master TFs that govern this process.

Conclusion

Enhancers are non-coding DNA elements that control cell-type-specific gene expression programs during development. These *cis*-regulatory elements facilitate the translation of

cellular signals into specific transcriptional outcomes through the binding of TFs. Thus, the binding of specific sets of TFs at enhancers can lead to the selective activation of target genes in a cell type- and context-dependent manner. Advances in sequencing technologies and computational approaches have allowed the identification of a large number of putative enhancers, enabling a better understanding of how specific TFs regulate development and how faulty regulation can lead to disease. Despite the wealth of available data, there are important caveats to consider when extrapolating these studies to a broad understanding of enhancer regulation. Single-cell technologies will be instrumental in defining the key mechanisms that lead to enhancer activation, changes in chromatin structure and DNA topology, and ultimately gene activation. Thus, additional insights from systematic studies will allow us to fully understand the complexity of developmental gene regulation and will open new avenues for identifying the genetic culprits that underpin developmental disorders such as congenital heart defects.

REFERENCES

Adachi, K., Kopp, W., Wu, G., Heising, S., Greber, B., Stehling, M., Araúzo-Bravo, M.J., Boerno, S.T., Timmermann, B., Vingron, M., et al. (2018). Esrrb Unlocks Silenced Enhancers for Reprogramming to Naive Pluripotency. Cell Stem Cell *23*, 266-275.e6.

Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nature Reviews Genetics *13*, 720–731.

Aguilo, F., Li, S., Balasubramaniyan, N., Sancho, A., Benko, S., Zhang, F., Vashisht, A., Rengasamy, M., Andino, B., Chen, C., et al. (2016). Deposition of 5-Methylcytosine on Enhancer RNAs Enables the Coactivator Function of PGC-1α. Cell Reports *14*, 479–492.

Ahmad, S.M., Busser, B.W., Huang, D., Cozart, E.J., Michaud, S., Zhu, X., Jeffries, N., Aboukhalil, A., Bulyk, M.L., Ovcharenko, I., et al. (2014). Machine learning classification of cell-specific cardiac enhancers uncovers developmental subnetworks regulating progenitor cell division and cell fate specification. Development *141*, 878–888.

Akerberg, B.N., Gu, F., VanDusen, N.J., Zhang, X., Dong, R., Li, K., Zhang, B., Zhou, B., Sethi, I., Ma, Q., et al. (2019). A reference map of murine cardiac transcription factor chromatin occupancy identifies dynamic and conserved enhancers. Nat Commun *10*, 4907.

Alvarez-Dominguez, J.R., Knoll, M., Gromatzky, A.A., and Lodish, H.F. (2017). The Super-Enhancer-Derived alncRNA-EC7/Bloodlinc Potentiates Red Blood Cell Development in trans. Cell Reports *19*, 2503–2514.

Ambrosino, C., Iwata, T., Scafoglio, C., Mallardo, M., Klein, R., and Nebreda, A.R. (2006). TEF-1 and C/EBPbeta are major p38alpha MAPK-regulated transcription factors in proliferating cardiomyocytes. The Biochemical Journal *396*, 163–172.

Andersson, R., and Sandelin, A. (2019). Determinants of enhancer and promoter activities of regulatory elements. Nat Rev Genet *21*, 71–87.

Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al. (2014). An atlas of active enhancers across human cell types and tissues. Nature *507*, 455–461.

Arai, A., Yamamoto, K., and Toyama, J. (1997). Murine cardiac progenitor cells require visceral embryonic endoderm and primitive streak for terminal differentiation. Developmental Dynamics *210*, 344–353.

Arnold, C.D., Gerlach, D., Stelzer, C., Boryn, L.M., Rath, M., and Stark, A. (2013a). Genome-Wide Quantitative Enhancer Activity Maps Identified by STARR-seq. Science *339*, 1074–1077.

Arnold, P., Schöler, A., Pachkov, M., Balwierz, P.J., Jørgensen, H., Stadler, M.B., Nimwegen, E.

van, and Schübeler, D. (2013b). Modeling of epigenome dynamics identifies transcription factors that mediate Polycomb targeting. Genome Research *23*, 60–73.

Arnold, P.R., Wells, A.D., and Li, X.C. (2019). Diversity and Emerging Roles of Enhancer RNA in Regulation of Gene Expression and Cell Fate. Frontiers Cell Dev Biology *7*, 377.

Atchison, M.L. (1988). Enhancers: Mechanisms of Action and Cell Specificity. Annual Review of Cell Biology *4*, 127–153.

Avruch, J., Zhou, D., Fitamant, J., Bardeesy, N., Mou, F., and Barrufet, L.R. (2012). Protein kinases of the Hippo pathway: Regulation and substrates. Seminars in Cell & Developmental Biology *23*, 770–784.

Azakie, A., LaMont, L., Fineman, J.R., and He, Y. (2005). Divergent transcriptional enhancer factor-1 regulates the cardiac troponin T promoter. American Journal of Physiology-Cell Physiology *289*, C1522–C1534.

Azofeifa, J.G., Allen, M.A., Hendrix, J.R., Read, T., Rubin, J.D., and Dowell, R.D. (2018). Enhancer RNA profiling predicts transcription factor activity. Genome Res *28*, 334–344.

Baillat, D., Hakimi, M.-A., Näär, A.M., Shilatifard, A., Cooch, N., and Shiekhattar, R. (2005). Integrator, a Multiprotein Mediator of Small Nuclear RNA Processing, Associates with the C-Terminal Repeat of RNA Polymerase II. Cell *123*, 265–276.

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299–308.

Barbieri, E., Trizzino, M., Welsh, S.A., Owens, T.A., Calabretta, B., Carroll, M., Sarma, K., and Gardini, A. (2018). Targeted Enhancer Activation by a Subunit of the Integrator Complex. Mol Cell *71*, 103-116.e7.

Barolo, S., and Posakony, J.W. (2002). Three habits of highly effective signaling pathways: principles of transcriptional control by developmental cell signaling. Gene Dev *16*, 1167–1181.

Barrera, L.A., Vedenko, A., Kurland, J.V., Rogers, J.M., Gisselbrecht, S.S., Rossin, E.J., Woodard, J., Mariani, L., Kock, K.H., Inukai, S., et al. (2016). Survey of variation in human transcription factors reveals prevalent DNA binding changes. Science *351*, 1450–1454.

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-Resolution Profiling of Histone Methylations in the Human Genome. Cell *129*, 823–837.

Basson, C.T., Bachinsky, D.R., Lin, R.C., Levi, T., Elkins, J.A., Soults, J., Grayzel, D., Kroumpouzou, E., Traill, T.A., Leblanc-Straceski, J., et al. (1997). Mutations in human cause limb and cardiac

malformation in Holt-Oram syndrome. Nature Genetics *15*, 30–35.

Ben-Levy, R., Hooper, S., Wilson, R., Paterson, H.F., and Marshall, C.J. (1998). Nuclear export of the stress-activated protein kinase p38 mediated by its substrate MAPKAP kinase-2. Current Biology *8*, 1049–1057.

Bentsen, M., Goymann, P., Schultheis, H., Klee, K., Petrova, A., Wiegandt, R., Fust, A., Preussner, J., Kuenne, C., Braun, T., et al. (2020). Beyond accessibility: ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. Biorxiv 869560.

Benveniste, D., Sonntag, H.-J., Sanguinetti, G., and Sproul, D. (2014). Transcription factor binding predicts histone modifications in human cell lines. Proceedings of the National Academy of Sciences *111*, 13367–13372.

Berkes, C.A., Bergstrom, D.A., Penn, B.H., Seaver, K.J., Knoepfler, P.S., and Tapscott, S.J. (2004). Pbx Marks Genes for Activation by MyoD Indicating a Role for a Homeodomain Protein in Establishing Myogenic Potential. Mol Cell *14*, 465–477.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Stamatoyannopoulos, J.A., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799–816.

Björnström, L., and Sjöberg, M. (2005). Mechanisms of Estrogen Receptor Signaling: Convergence of Genomic and Nongenomic Actions on Target Genes. Mol Endocrinol *19*, 833–842.

Blue, G.M., Kirk, E.P., Giannoulatou, E., Sholler, G.F., Dunwoodie, S.L., Harvey, R.P., and Winlaw, D.S. (2017). Advances in the Genetics of Congenital Heart Disease: A Clinician's Guide. J Am Coll Cardiol *69*, 859–870.

Bondue, A., Lapouge, G., Paulissen, C., Semeraro, C., Iacovino, M., Kyba, M., and Blanpain, C. (2008). Mesp1 acts as a master regulator of multipotent cardiovascular progenitor specification. Cell Stem Cell *3*, 69–84.

Bonev, B., Cohen, N.M., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.-P., Tanay, A., et al. (2017). Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell *171*, 557-572.e24.

Bonn, S., Zinzen, R.P., Girardot, C., Gustafson, E.H., Perez-Gonzalez, A., Delhomme, N., Ghavi-Helm, Y., Wilczyński, B., Riddell, A., and Furlong, E.E.M. (2012). Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. Nature Genetics *44*, 148–156.

Bose, D.A., Donahue, G., Reinberg, D., Shiekhattar, R., Bonasio, R., and Berger, S.L. (2017). RNA Binding to CBP Stimulates Histone Acetylation and Transcription. Cell *168*, 135-149.e22.

Bossard, P., and Zaret, K.S. (1998). GATA transcription factors as potentiators of gut endoderm differentiation. Development *125*, 4909–4917.

Boyle, A.P., Davis, S., Shulha, H.P., Meltzer, P., Margulies, E.H., Weng, Z., Furey, T.S., and Crawford, G.E. (2008). High-resolution mapping and characterization of open chromatin across the genome. Cell *132*, 311–322.

Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A., and Cedar, H. (1994). Sp1 elements protect a CpG island from de novo methylation. Nature *371*, 435–438.

Bruneau, B.G. (2008). The developmental genetics of congenital heart disease. Nature *451*, 943–948.

Bruneau, B.G. (2013). Signaling and Transcriptional Networks in Heart Development and Regeneration. Cold Spring Harbor Perspectives in Biology *5*, a008292–a008292.

Brunelle, M., Markovits, A.N., Rodrigue, S., Lupien, M., Jacques, P.-É., and Gévry, N. (2015). The histone variant H2A.Z is an important regulator of enhancer activity. Nucleic Acids Research *43*, 9742–9756.

Bu, L., Jiang, X., Martin-Puig, S., Caron, L., Zhu, S., Shao, Y., Roberts, D.J., Huang, P.L., Domian, I.J., and Chien, K.R. (2009). Human ISL1 heart progenitors generate diverse multipotent cardiovascular cell lineages. Nature *460*, 113–117.

Buckingham, M., Meilhac, S., and Zaffran, S. (2005). Building the mammalian heart from two sources of myocardial cells. Nature Reviews Genetics *6*, 826–835.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature Methods *10*, 1213–1218.

Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015a). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide (John Wiley & Sons, Inc.).

Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., and Greenleaf, W.J. (2015b). Single-cell chromatin accessibility reveals principles of regulatory variation. Nature *523*, 486–490.

Buikema, J.W., Mady, A.S., Mittal, N.V., Atmanli, A., Caron, L., Doevendans, P.A., Sluijter, J.P.G., and Domian, I.J. (2013). Wnt/β-catenin signaling directs the regional expansion of first and

second heart field-derived ventricular cardiomyocytes. Dev Camb Engl *140*, 4165–4176.

Bulger, M., and Groudine, M. (2010). Enhancers: The abundance and function of regulatory sequences beyond promoters. Developmental Biology *339*, 250–257.

Bulger, M., and Groudine, M. (2011). Functional and Mechanistic Diversity of Distal Transcription Enhancers. Cell *144*, 327–339.

Bürglin, T.R. (1991). The TEA domain: A novel, highly conserved DNA-binding motif. Cell *66*, 11–12.

Canver, M.C., Smith, E.C., Sher, F., Pinello, L., Sanjana, N.E., Shalem, O., Chen, D.D., Schupp, P.G., Vinjamur, D.S., Garcia, S.P., et al. (2015). BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. Nature *527*, 192–197.

Cao, Q., Anyansi, C., Hu, X., Xu, L., Xiong, L., Tang, W., Mok, M.T.S., Cheng, C., Fan, X., Gerstein, M., et al. (2017). Reconstruction of enhancer–target networks in 935 samples of human primary cells, tissues and cell lines. Nat Genet *49*, 1428–1436.

Cernilogar, F.M., Hasenöder, S., Wang, Z., Scheibner, K., Burtscher, I., Sterr, M., Smialowski, P., Groh, S., Evenroed, I.M., Gilfillan, G.D., et al. (2019). Pre-marked chromatin and transcription factor co-binding shape the pioneering activity of Foxa2. Nucleic Acids Res *47*, 9069–9086.

Chan, S.S.-K., and Kyba, M. (2013). What is a Master Regulator? J Stem Cell Res Ther *03*.

Chang, C.-P., and Bruneau, B.G. (2012). Epigenetics and Cardiovascular Development. Annual Review of Physiology *74*, 41–68.

Chang, L., Azzolin, L., Biagio, D.D., Zanconato, F., Battilana, G., Xiccato, R.L., Aragona, M., Giulitti, S., Panciera, T., Gandin, A., et al. (2018). The SWI/SNF complex is a mechanoregulated inhibitor of YAP and TAZ. Nature *563*, 265–269.

Chen, F.X., Xie, P., Collings, C.K., Cao, K., Aoi, Y., Marshall, S.A., Rendleman, E.J., Ugarenko, M., Ozark, P.A., Zhang, A., et al. (2017). PAF1 regulation of promoter-proximal pause release via enhancer activation. Sci New York N Y *357*, 1294–1298.

Chen, Q., Zhang, H., Liu, Y., Adams, S., Eilken, H., Stehling, M., Corada, M., Dejana, E., Zhou, B., and Adams, R.H. (2016). Endothelial cells are progenitors of cardiac pericytes and vascular smooth muscle cells. Nat Commun *7*, 12422.

Chen, S.N., Gurha, P., Lombardi, R., Ruggiero, A., Willerson, J.T., and Marian, A.J. (2014). The Hippo Pathway Is Activated and Is a Causal Mechanism for Adipogenesis in Arrhythmogenic Cardiomyopathy. Circ Res *114*, 454–468.

Chen, Z., Friedrich, G.A., and Soriano, P. (1994). Transcriptional enhancer factor 1 disruption by a retroviral gene trap leads to heart defects and embryonic lethality in mice. Gene Dev *8*, 2293–2301.

Cirillo, L.A., Lin, F.R., Cuesta, I., Friedman, D., Jarnik, M., and Zaret, K.S. (2002). Opening of Compacted Chromatin by Early Developmental Transcription Factors HNF3 (FoxA) and GATA-4. Mol Cell *9*, 279–289.

Clark, K.L., Halay, E.D., Lai, E., and Burley, S.K. (1993). Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. Nature *364*, 412–420.

Cong, F., Schweizer, L., Chamorro, M., and Varmus, H. (2003). Requirement for a Nuclear Function of β-Catenin in Wnt Signaling. Mol Cell Biol *23*, 8462–8470.

Consortium, T.M.E., Yue, F., Cheng, Y., Breschi, A., Vierstra, J., Wu, W., Ryba, T., Sandstrom, R., Ma, Z., Davis, C., et al. (2014). A comparative encyclopedia of DNA elements in the mouse genome. Nature *515*, 355–364.

Corces, M.R., Buenrostro, J.D., Wu, B., Greenside, P.G., Chan, S.M., Koenig, J.L., Snyder, M.P., Pritchard, J.K., Kundaje, A., Greenleaf, W.J., et al. (2016). Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nature Genetics *48*, 1193–1203.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nature Methods *14*, 959–962.

Corradin, O., Saiakhova, A., Akhtar-Zaidi, B., Myeroff, L., Willis, J., dot}lari, R.C.-S., Lupien, M., Markowitz, S., and Scacheri, P.C. (2013). Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits. Genome Res *24*, 1–13.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences *107*, 21931–21936.

Cusanovich, D.A., Daza, R., Adey, A., Pliner, H.A., Christiansen, L., Gunderson, K.L., Steemers, F.J., Trapnell, C., and Shendure, J. (2015). Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. Science *348*, 910–914.

Dalvai, M., Fleury, L., Bellucci, L., Kocanova, S., and Bystricky, K. (2013). TIP48/Reptin and H2A.Z requirement for initiating chromatin remodeling in estrogen-activated transcription. Plos Genet *9*, e1003387.

Dancy, B.M., and Cole, P.A. (2014). Protein Lysine Acetylation by p300/CBP. Chemical Reviews *115*, 2419–2452.

Davidson, I., Xiao, J.H., Rosales, R., Staub, A., and Chambon, P. (1988). The HeLa cell protein TEF-1 binds specifically and cooperatively to two SV40 enhancer motifs of unrelated sequence. Cell *54*, 931–942.

Deaton, A.M., Gómez-Rodríguez, M., Mieczkowski, J., Tolstorukov, M.Y., Kundu, S., Sadreyev, R.I., Jansen, L.E., and Kingston, R.E. (2016). Enhancer regions show high histone H3.3 turnover that changes during differentiation. ELife *5*, e15316.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing Chromosome Conformation. Science *295*, 1306–1311.

Del Re, D.P., Matsuda, T., Zhai, P., Maejima, Y., Jain, M.R., Liu, T., Li, H., Hsu, C.-P., and Sadoshima, J. (2014). Mst1 Promotes Cardiac Myocyte Apoptosis through Phosphorylation and Inhibition of Bcl-xL. Mol Cell *54*, 639–650.

Deplancke, B., Alpern, D., and Gardeux, V. (2016). The Genetics of Transcription Factor DNA Binding Variation. Cell *166*, 538–554.

Devine, W.P., Wythe, J.D., George, M., Koshiba-Takeuchi, K., and Bruneau, B.G. (2014). Early patterning and specification of cardiac progenitors in gastrulating mesoderm. ELife *3*, 508.

Diao, Y., Li, B., Meng, Z., Jung, I., Lee, A.Y., Dixon, J., Maliskova, L., Guan, K.-L., Shen, Y., and Ren, B. (2016). A new class of temporarily phenotypic enhancers identified by CRISPR/Cas9-mediated genetic screening. Genome Research *26*, 397–405.

Dickel, D.E., Barozzi, I., Zhu, Y., Fukuda-Yuzawa, Y., Osterwalder, M., Mannion, B.J., May, D., Spurrell, C.H., Plajzer-Frick, I., Pickle, C.S., et al. (2016). Genome-wide compendium and functional assessment of *in vivo* heart enhancers. Nature Communications *7*, 12923.

Dickinson, M.E., Flenniken, A.M., Ji, X., Teboul, L., Wong, M.D., White, J.K., Meehan, T.F., Weninger, W.J., Westerberg, H., Adissu, H., et al. (2016). High-throughput discovery of novel developmental phenotypes. Nature *537*, 508–514.

Diez-Cuñado, M., Wei, K., Bushway, P.J., Maurya, M.R., Perera, R., Subramaniam, S., Ruiz-Lozano, P., and Mercola, M. (2018). miRNAs that Induce Human Cardiomyocyte Proliferation Converge on the Hippo Pathway. Cell Reports *23*, 2168–2174.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331–336.

Donaghey, J., Thakurela, S., Charlton, J., Chen, J.S., Smith, Z.D., Gu, H., Pop, R., Clement, K., Stamenova, E.K., Karnik, R., et al. (2018). Genetic determinants and epigenetic effects of pioneer-factor occupancy. Nat Genet *50*, 250–258.

Dong, J., Feldmann, G., Huang, J., Wu, S., Zhang, N., Comerford, S.A., Gayyed, M.F., Anders, R.A., Maitra, A., and Pan, D. (2007). Elucidation of a Universal Size-Control Mechanism in Drosophila and Mammals. Cell *130*, 1120–1133.

Dorighi, K.M., Swigut, T., Henriques, T., Bhanu, N.V., Scruggs, B.S., Nady, N., Still, C.D., Garcia, B.A., Adelman, K., and Wysocka, J. (2017). Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation. Mol Cell *66*, 568-576.e4.

Dowen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schujiers, J., Lee, T.I., Zhao, K., et al. (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. Cell *159*, 374–387.

Dukler, N., Gulko, B., Huang, Y.-F., and Siepel, A. (2016). Is a super-enhancer greater than the sum of its parts? Nat Genet *49*, 2–3.

Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J., Kaul, R., et al. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Ecco, G., Cassano, M., Kauzlaric, A., Duc, J., Coluccio, A., Offner, S., Imbeault, M., Rowe, H.M., Turelli, P., and Trono, D. (2016). Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues. Dev Cell *36*, 611–623.

Ege, N., Dowbaj, A.M., Jiang, M., Howell, M., Hooper, S., Foster, C., Jenkins, R.P., and Sahai, E. (2018). Quantitative Analysis Reveals that Actin and Src-Family Kinases Regulate Nuclear YAP1 and Its Export. Cell Syst *6*, 692-708.e13.

Elgin, S.C.R. (1981). DNAase I-hypersensitive sites of chromatin. Cell *27*, 413–415.

Elosegui-Artola, A., Andreu, I., Beedle, A.E.M., Lezamiz, A., Uroz, M., Kosmalska, A.J., Oria, R., Kechagia, J.Z., Rico-Lastres, P., Roux, A.-L.L., et al. (2017). Force Triggers YAP Nuclear Entry by Regulating Transport across Nuclear Pores. Cell *171*, 1397-1410.e14.

Elrod, N.D., Henriques, T., Huang, K.-L., Tatomer, D.C., Wilusz, J.E., Wagner, E.J., and Adelman, K. (2019). The Integrator Complex Attenuates Promoter-Proximal Transcription at Protein-

Coding Genes. Mol Cell *76*, 738-752.e7.

Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., et al. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. Nature *473*, 43–49.

Ernst, J., Melnikov, A., Zhang, X., Wang, L., Rogov, P., Mikkelsen, T.S., and Kellis, M. (2016). Genome-scale high-resolution mapping of activating and repressive nucleotides in regulatory regions. Nat Biotechnol *34*, 1180–1190.

Farley, E.K., Olson, K.M., Zhang, W., Brandt, A.J., Rokhsar, D.S., and Levine, M.S. (2015). Suboptimization of developmental enhancers. Sci New York N Y *350*, 325–328.

Farley, E.K., Olson, K.M., Zhang, W., Rokhsar, D.S., and Levine, M.S. (2016). Syntax compensates for poor binding sites to encode tissue specificity of developmental enhancers. P Natl Acad Sci Usa *113*, 6508–6513.

Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Stein, T.I., Rosen, N., Kohn, A., Twik, M., Safran, M., et al. (2017). GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database J Biological Databases Curation *2017*, bax028.

Flores, M.A., and Ovcharenko, I. (2018). Enhancer reprogramming in mammalian genomes. Bmc Bioinformatics *19*, 316.

Freire-Pritchett, P., Schoenfelder, S., Várnai, C., Wingett, S.W., Cairns, J., Collier, A.J., García-Vílchez, R., Furlan-Magaril, M., Osborne, C.S., Fraser, P., et al. (2017). Global reorganisation of cis-regulatory units upon lineage commitment of human embryonic stem cells. Elife *6*, e21926.

Friman, E.T., Deluz, C., Meireles-Filho, A.C., Govindan, S., Gardeux, V., Deplancke, B., and Suter, D.M. (2019). Dynamic regulation of chromatin accessibility by pluripotency transcription factors across the cell cycle. Elife *8*, e50087.

Fu, V., Plouffe, S.W., and Guan, K.-L. (2017). The Hippo pathway in organ development, homeostasis, and regeneration. Curr Opin Cell Biol *49*, 99–107.

Fuglerud, B.M., Ledsaak, M., Rogne, M., Eskeland, R., and Gabrielsen, O.S. (2018). The pioneer factor activity of c-Myb involves recruitment of p300 and induction of histone acetylation followed by acetylation-induced chromatin dissociation. Epigenet Chromatin *11*, 35.

Fujiki, K., Shinoda, A., Kano, F., Sato, R., Shirahige, K., and Murata, M. (2013). PPARγ-induced PARylation promotes local DNA demethylation by production of 5-hydroxymethylcytosine. Nature Communications *4*, 2262.

Fulco, C.P., Munschauer, M., Anyoha, R., Munson, G., Grossman, S.R., Perez, E.M., Kane, M.,

Cleary, B., Lander, E.S., and Engreitz, J.M. (2016). Systematic mapping of functional enhancer-promoter connections with CRISPR interference. Science *354*, 769–773.

Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H., et al. (2009). An oestrogen-receptor-α-bound human chromatin interactome. Nature *462*, 58–64.

Galdos, F.X., Guo, Y., Paige, S.L., VanDusen, N.J., Wu, S.M., and Pu, W.T. (2017). Cardiac Regeneration: Lessons From Development. Circ Res *120*, 941–959.

Galli, G.G., Carrara, M., Yuan, W.-C., Valdes-Quezada, C., Gurung, B., Pepe-Mooney, B., Zhang, T., Geeven, G., Gray, N.S., Laat, W. de, et al. (2015). YAP Drives Growth by Controlling Transcriptional Pause Release from Dynamic Enhancers. Molecular Cell *60*, 328–337.

Gao, R., Liang, X., Cheedipudi, S., Cordero, J., Jiang, X., Zhang, Q., Caputo, L., Günther, S., Kuenne, C., Ren, Y., et al. (2019). Pioneering function of Isl1 in the epigenetic control of cardiomyocyte cell fate. Cell Res *29*, 486–501.

Garcia, M.F., Moore, C.D., Schulz, K.N., Alberto, O., Donague, G., Harrison, M.M., Zhu, H., and Zaret, K.S. (2019). Structural Features of Transcription Factors Associating with Nucleosome Binding. Mol Cell *75*, 921-932.e6.

Garton, M., Najafabadi, H.S., Schmitges, F.W., Radovani, E., Hughes, T.R., and Kim, P.M. (2015). A structural approach reveals how neighbouring C2H2 zinc fingers influence DNA binding specificity. Nucleic Acids Res *43*, 9147–9157.

Gasperini, M., Hill, A.J., McFaline-Figueroa, J.L., Martin, B., Kim, S., Zhang, M.D., Jackson, D., Leith, A., Schreiber, J., Noble, W.S., et al. (2019). A Genome-wide Framework for Mapping Gene Regulation via Cellular Genetic Screens. Cell *176*, 377-390.e19.

Gasperini, M., Tome, J.M., and Shendure, J. (2020). Towards a comprehensive catalogue of validated and target-linked human enhancers. Nat Rev Genetics 1–19.

Gelb, B.D., and Chung, W.K. (2014). Complex Genetics and the Etiology of Human Congenital Heart Disease. Cold Spring Harbor Perspectives in Medicine *4*, a013953–a013953.

Ghavi-Helm, Y., Klein, F.A., Pakozdi, T., Ciglar, L., Noordermeer, D., Huber, W., and Furlong, E.E.M. (2014). Enhancer loops appear stable during development and are associated with paused polymerase. Nature *512*, 96–100.

Gise, A. von, Lin, Z., Schlegelmilch, K., Honor, L.B., Pan, G.M., Buck, J.N., Ma, Q., Ishiwata, T., Zhou, B., Camargo, F.D., et al. (2012). YAP1, the nuclear target of Hippo signaling, stimulates heart growth through cardiomyocyte proliferation but not hypertrophy. Proceedings of the National Academy of Sciences *109*, 2394–2399.

Goldberg, A.D., Banaszynski, L.A., Noh, K.-M., Lewis, P.W., Elsaesser, S.J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X., et al. (2010). Distinct Factors Control Histone Variant H3.3 Localization at Specific Genomic Regions. Cell *140*, 678–691.

Gosselin, D., Link, V.M., Romanoski, C.E., Fonseca, G.J., Eichenfield, D.Z., Spann, N.J., Stender, J.D., Chun, H.B., Garner, H., Geissmann, F., et al. (2014). Environment drives selection and function of enhancers controlling tissue-specific macrophage identities. Cell *159*, 1327–1340.

Goulev, Y., Fauny, J.D., Gonzalez-Marti, B., Flagiello, D., Silber, J., and Zider, A. (2008). SCALLOPED Interacts with YORKIE, the Nuclear Effector of the Hippo Tumor-Suppressor Pathway in Drosophila. Current Biology *18*, 435–441.

Gusmao, E.G., Dieterich, C., Zenke, M., and Costa, I.G. (2014). Detection of active transcription factor binding sites with the combination of DNase hypersensitivity and histone modifications. Bioinform Oxf Engl *30*, 3143–3151.

Hafez, D., Karabacak, A., Krueger, S., Hwang, Y.-C., Wang, L.-S., Zinzen, R.P., and Ohler, U. (2017). McEnhancer: predicting gene expression via semi-supervised assignment of enhancers to target genes. Genome Biol *18*, 199.

Hah, N., Murakami, S., Nagari, A., Danko, C.G., and Kraus, W.L. (2013). Enhancer transcripts mark active estrogen receptor binding sites. Genome Res *23*, 1210–1223.

Hait, T.A., Amar, D., Shamir, R., and Elkon, R. (2018). FOCS: a novel method for analyzing enhancer and gene activity patterns infers an extensive enhancer–promoter map. Genome Biol *19*, 56.

Halfon, M.S. (2019). Studying Transcriptional Enhancers: The Founder Fallacy, Validation Creep, and Other Biases. Trends Genet *35*, 93–103.

Hariprakash, J.M., and Ferrari, F. (2019). Computational Biology Solutions to Identify Enhancers-target Gene Pairs. Comput Struct Biotechnology J *17*, 821–831.

Harvey, K.F., Zhang, X., and Thomas, D.M. (2013). The Hippo pathway and human cancer. Nature Reviews Cancer *13*, 246–257.

Hay, D., Hughes, J.R., Babbs, C., Davies, J.O.J., Graham, B.J., Hanssen, L.L.P., Kassouf, M.T., Oudelaar, A.M., Sharpe, J.A., Suciu, M.C., et al. (2016). Genetic dissection of the α-globin super-enhancer in vivo. Nat Genet *48*, 895–903.

He, A., Kong, S.W., Ma, Q., and Pu, W.T. (2011). Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. Proc National Acad Sci *108*, 5632–5637.

He, B., Chen, C., Teng, L., and Tan, K. (2014). Global view of enhancer–promoter interactome in human cells. Proc National Acad Sci *111*, E2191–E2199.

Heallen, T., Zhang, M., Wang, J., Bonilla-Claudio, M., Klysik, E., Johnson, R.L., and Martin, J.F. (2011). Hippo Pathway Inhibits Wnt Signaling to Restrain Cardiomyocyte Proliferation and Heart Size. Science *332*, 458–461.

Heallen, T., Morikawa, Y., Leach, J., Tao, G., Willerson, J.T., Johnson, R.L., and Martin, J.F. (2013). Hippo signaling impedes adult heart regeneration. Development *140*, 4683–4690.

Henriques, T., Scruggs, B.S., Inouye, M.O., Muse, G.W., Williams, L.H., Burkholder, A.B., Lavender, C.A., Fargo, D.C., and Adelman, K. (2018). Widespread transcriptional pausing and elongation control at enhancers. Gene Dev *32*, 26–41.

Herz, H.-M., Mohan, M., Garruss, A.S., Liang, K., Takahashi, Y.-H., Mickey, K., Voets, O., Verrijzer, C.P., and Shilatifard, A. (2012). Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. Gene Dev *26*, 2604–2620.

Hilton, I.B., D'Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nature Biotechnology *33*, 510–517.

Hiroi, Y., Kudoh, S., Monzen, K., Ikeda, Y., Yazaki, Y., Nagai, R., and Komuro, I. (2001). Tbx5 associates with Nkx2-5 and synergistically promotes cardiomyocyte differentiation. Nat Genet *28*, 276–280.

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-Enhancers in the Control of Cell Identity and Disease. Cell *155*, 934–947.

Hnisz, D., Day, D.S., and Young, R.A. (2016). Insulated Neighborhoods: Structural and Functional Units of Mammalian Gene Control. Cell *167*, 1188–1200.

Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K., and Sharp, P.A. (2017). A Phase Separation Model for Transcriptional Control. Cell *169*, 13–23.

Holm, H., Gudbjartsson, D.F., Arnar, D.O., Thorleifsson, G., Thorgeirsson, G., Stefansdottir, H., Gudjonsson, S.A., Jonasdottir, A., Mathiesen, E.B., Njølstad, I., et al. (2010). Several common variants modulate heart rate, PR interval and QRS duration. Nature Genetics *42*, 117–122.

Holt, M., and Oram, S. (1960). FAMILIAL HEART DISEASE WITH SKELETAL MALFORMATIONS. Heart *22*, 236–242.

Homsy, J., Zaidi, S., Shen, Y., Ware, J.S., Samocha, K.E., Karczewski, K.J., DePalma, S.R., McKean, D., Wakimoto, H., Gorham, J., et al. (2015). De novo mutations in congenital heart disease with

neurodevelopmental and other congenital anomalies. Sci New York N Y *350*, 1262–1266.

Hossain, Z., Ali, S.M., Ko, H.L., Xu, J., Ng, C.P., Guo, K., Qi, Z., Ponniah, S., Hong, W., and Hunziker, W. (2007). Glomerulocystic kidney disease in mice with a targeted inactivation of Wwtr1. Proc National Acad Sci *104*, 1631–1636.

Hsieh, C.-L., Fei, T., Chen, Y., Li, T., Gao, Y., Wang, X., Sun, T., Sweeney, C.J., Lee, G.-S.M., Chen, S., et al. (2014). Enhancer RNAs participate in androgen receptor-driven looping that selectively enhances gene activation. Proc National Acad Sci *111*, 7319–7324.

Huang, J., Wu, S., Barrera, J., Matthews, K., and Pan, D. (2005). The Hippo signaling pathway coordinately regulates cell proliferation and apoptosis by inactivating Yorkie, the Drosophila Homolog of YAP. Cell *122*, 421–434.

Huertas, J., MacCarthy, C.M., Schöler, H.R., and Cojocaru, V. (2020). Nucleosomal DNA Dynamics Mediate Oct4 Pioneer Factor Binding. Biophys J.

Ieda, M., Fu, J.-D., Delgado-Olguin, P., Vedantham, V., Hayashi, Y., Bruneau, B.G., and Srivastava, D. (2010). Direct Reprogramming of Fibroblasts into Functional Cardiomyocytes by Defined Factors. Cell *142*, 375–386.

Ikeda, S., Mizushima, W., Sciarretta, S., Abdellatif, M., Zhai, P., Mukai, R., Fefelova, N., Oka, S., Nakamura, M., Re, D.P.D., et al. (2019). Hippo Deficiency Leads to Cardiac Dysfunction Accompanied by Cardiomyocyte Dedifferentiation During Pressure Overload. Circ Res *124*, 292–305.

Inukai, S., Kock, K.H., and Bulyk, M.L. (2017). Transcription factor–DNA binding: beyond binding site motifs. Curr Opin Genet Dev *43*, 110–119.

Ivey, M.J., and Tallquist, M.D. (2016). Defining the Cardiac Fibroblast. Circ J *80*, 2269–2276.

Iwafuchi-Doi, M. (2018). The mechanistic basis for chromatin regulation by pioneer transcription factors. Wiley Interdiscip Rev Syst Biology Medicine *11*, e1427.

Iwafuchi-Doi, M., Donahue, G., Kakumanu, A., Watts, J.A., Mahony, S., Pugh, B.F., Lee, D., Kaestner, K.H., and Zaret, K.S. (2016). The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. Mol Cell *62*, 79–91.

Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol *3*, 318–356.

Jia, G., Preussner, J., Chen, X., Guenther, S., Yuan, X., Yekelchyk, M., Kuenne, C., Looso, M., Zhou, Y., Teichmann, S., et al. (2018). Single cell RNA-seq and ATAC-seq analysis of cardiac

progenitor cell transition states and lineage settlement. Nat Commun *9*, 4877.

Jiang, Y., Qian, F., Bai, X., Liu, Y., Wang, Q., Ai, B., Han, X., Shi, S., Zhang, J., Li, X., et al. (2018). SEdb: a comprehensive human super-enhancer database. Nucleic Acids Res *47*, D235–D243.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant–containing nucleosomes mark "nucleosome-free regions" of active promoters and other regulatory regions. Nature Genetics *41*, 941–945.

Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.-A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature *503*, 290–294.

Jin, S.C., Homsy, J., Zaidi, S., Lu, Q., Morton, S., DePalma, S.R., Zeng, X., Qi, H., Chang, W., Sierant, M.C., et al. (2017). Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. Nat Genet *49*, 1593–1601.

Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., et al. (2013). DNA-binding specificities of human transcription factors. Cell *152*, 327–339.

Jolma, A., Yin, Y., Nitta, K.R., Dave, K., Popov, A., Taipale, M., Enge, M., Kivioja, T., Morgunova, E., and Taipale, J. (2015). DNA-dependent formation of transcription factor pairs alters their binding specificity. Nature *527*, 384–388.

Jozwik, K.M., Chernukhin, I., Serandour, A.A., Nagarajan, S., and Carroll, J.S. (2016). FOXA1 Directs H3K4 Monomethylation at Enhancers via Recruitment of the Methyltransferase MLL3. Cell Reports *17*, 2715–2723.

Junion, G., Spivakov, M., Girardot, C., Braun, M., Gustafson, E.H., Birney, E., and Furlong, E.E.M. (2012). A Transcription Factor Collective Defines Cardiac Cell Fate and Reflects Lineage History. Cell *148*, 473–486.

Kaikkonen, M.U., Spann, N.J., Heinz, S., Romanoski, C.E., Allison, K.A., Stender, J.D., Chun, H.B., Tough, D.F., Prinjha, R.K., Benner, C., et al. (2013). Remodeling of the Enhancer Landscape during Macrophage Activation Is Coupled to Enhancer Transcription. Molecular Cell *51*, 310–325.

Kaneko, K.J., Kohn, M.J., Liu, C., and DePamphilis, M.L. (2007). Transcription factor TEAD2 is involved in neural tube closure. Genesis *45*, 577–587.

Kanno, T., Kanno, Y., LeRoy, G., Campos, E., Sun, H.-W., Brooks, S.R., Vahedi, G., Heightman, T.D., Garcia, B.A., Reinberg, D., et al. (2014). BRD4 assists elongation of both coding and enhancer RNAs by interacting with acetylated histones. Nat Struct Mol Biol *21*, 1047–1057.

Karagianni, P., Moulos, P., Schmidt, D., Odom, D.T., and Talianidis, I. (2020). Bookmarking by Non-pioneer Transcription Factors during Liver Development Establishes Competence for Future Gene Activation. Cell Reports *30*, 1319-1328.e6.

Karbassi, E., Fenix, A., Marchiano, S., Muraoka, N., Nakamura, K., Yang, X., and Murry, C.E. (2020). Cardiomyocyte maturation: advances in knowledge and implications for regenerative medicine. Nat Rev Cardiol 1–19.

Kerkela, R., Kockeritz, L., Macaulay, K., Zhou, J., Doble, B.W., Beahm, C., Greytak, S., Woulfe, K., Trivedi, C.M., Woodgett, J.R., et al. (2008). Deletion of GSK-3beta in mice leads to hypertrophic cardiomyopathy secondary to cardiomyoblast hyperproliferation. J Clin Investigation *118*, 3609–3618.

Kerppola, T.K., and Curran, T. (1991). Fos-Jun heterodimers and jun homodimers bend DNA in opposite orientations: Implications for transcription factor cooperativity. Cell *66*, 317–326.

Khan, A., and Zhang, X. (2015). dbSUPER: a database of super-enhancers in mouse and human genome. Nucleic Acids Res *44*, D164-71.

Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., et al. (2010). Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. Nature *471*, 480–485.

Kim, H.S., Tan, Y., Ma, W., Merkurjev, D., Destici, E., Ma, Q., Suter, T., Ohgi, K., Friedman, M., Skowronska-Krawczyk, D., et al. (2018a). Pluripotency factors functionally premark cell-type-restricted enhancers in ES cells. Nature *556*, 510–514.

Kim, Y.J., Xie, P., Cao, L., Zhang, M.Q., and Kim, T.H. (2018b). Global transcriptional activity dynamics reveal functional enhancer RNAs. Genome Res *28*, 1799–1811.

King, H.W., and Klose, R.J. (2017). The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. Elife *6*, e22631.

Klann, T.S., Black, J.B., Chellappan, M., Safi, A., Song, L., Hilton, I.B., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2017). CRISPR-Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. Nat Biotechnol *35*, 561–568.

Klein, J.C., Chen, W., Gasperini, M., and Shendure, J. (2017). Identifying Novel Enhancer Elements with CRISPR-Based Screens. Acs Chem Biol *13*, 326–332.

Klemm, S.L., Shipony, Z., and Greenleaf, W.J. (2019). Chromatin accessibility and the regulatory epigenome. Nat Rev Genet *20*, 207–220.

Koche, R.P., Smith, Z.D., Adli, M., Gu, H., Ku, M., Gnirke, A., Bernstein, B.E., and Meissner, A. (2011). Reprogramming factor expression initiates widespread targeted chromatin remodeling. Cell Stem Cell *8*, 96–105.

Korkmaz, G., Lopes, R., Ugalde, A.P., Nevedomskaya, E., Han, R., Myacheva, K., Zwart, W., Elkon, R., and Agami, R. (2016). Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. Nature Biotechnology *34*, 192–198.

Kouno, T., Moody, J., Kwon, A.T.-J., Shibayama, Y., Kato, S., Huang, Y., Böttcher, M., Motakis, E., Mendez, M., Severin, J., et al. (2019). C1 CAGE detects transcription start sites and enhancer activity at single-cell resolution. Nat Commun *10*, 360.

Kowalczyk, M.S., Hughes, J.R., Garrick, D., Lynch, M.D., Sharpe, J.A., Sloane-Stanley, J.A., McGowan, S.J., De Gobbi, M., Hosseini, M., Vernimmen, D., et al. (2012). Intragenic Enhancers Act as Alternative Promoters. Mol Cell *45*, 447–458.

Kragesteen, B.K., Spielmann, M., Paliou, C., Heinrich, V., Schöpflin, R., Esposito, A., Annunziatella, C., Bianco, S., Chiariello, A.M., Jerković, I., et al. (2018). Dynamic 3D chromatin architecture contributes to enhancer specificity and limb morphogenesis. Nat Genet *50*, 1463–1473.

Kraushaar, D.C., Jin, W., Maunakea, A., Abraham, B., Ha, M., and Zhao, K. (2013). Genome-wide incorporation dynamics reveal distinct categories of turnover for the histone variant H3.3. Genome Biol *14*, R121.

Kribelbauer, J.F., Rastogi, C., Bussemaker, H.J., and Mann, R.S. (2019). Low-Affinity Binding Sites and the Transcription Factor Specificity Paradox in Eukaryotes. Annu Rev Cell Dev Bi *35*, 357–379.

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M.J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–330.

Laat, W. de, and Duboule, D. (2013). Topology of mammalian developmental enhancers and their regulatory landscapes. Nature *502*, 499–506.

Lai, F., Orom, U. a, Cesaroni, M., Beringer, M., Taatjes, D.J., Blobel, G.A., and Shiekhattar, R. (2013). Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature *494*, 497–501.

Lai, F., Gardini, A., Zhang, A., and Shiekhattar, R. (2015). Integrator mediates the biogenesis of enhancer RNAs. Nature *525*, 399–403.

Lai, J.K.H., Collins, M.M., Uribe, V., Jiménez-Amilburu, V., Günther, S., Maischein, H.-M., and

Stainier, D.Y.R. (2018). The Hippo pathway effector Wwtr1 regulates cardiac wall maturation in zebrafish. Dev Camb Engl *145*, dev159210.

Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., and Weirauch, M.T. (2018). The Human Transcription Factors. Cell *172*, 650–665.

Lange, F.J. de, Moorman, A.F.M., Anderson, R.H., Männer, J., Soufan, A.T., Vries, C. de G., Schneider, M.D., Webb, S., Hoff, M.J.B. van den, and Christoffels, V.M. (2004). Lineage and Morphogenetic Analysis of the Cardiac Valves. Circ Res *95*, 645–654.

Leach, J.P., Heallen, T., Zhang, M., Rahmani, M., Morikawa, Y., Hill, M.C., Segura, A., Willerson, J.T., and Martin, J.F. (2017). Hippo pathway deficiency reverses systolic heart failure after infarction. Nature *550*, 260–264.

Lebrecht, D., Foehr, M., Smith, E., Lopes, F.J.P., Vanario-Alonso, C.E., Reinitz, J., Burz, D.S., and Hanes, S.D. (2005). Bicoid cooperative DNA binding is critical for embryonic patterning in Drosophila. Proc National Acad Sci *102*, 13176–13181.

Lescroart, F., Chabab, S., Lin, X., Rulands, S., Paulissen, C., Rodolosse, A., Auer, H., Achouri, Y., Dubois, C., Bondue, A., et al. (2014). Early lineage restriction in temporally distinct populations of Mesp1 progenitors during mammalian heart development. Nat Cell Biol *16*, 829–840.

Lescroart, F., Wang, X., Lin, X., Swedlund, B., Gargouri, S., Sànchez-Dànes, A., Moignard, V., Dubois, C., Paulissen, C., Kinston, S., et al. (2018). Defining the earliest step of cardiovascular lineage segregation by single-cell RNA-seq. Sci New York N Y *359*, eaao4174.

Lettice, L.A., Heaney, S.J.H., Purdie, L.A., Li, L., Beer, P. de, Oostra, B.A., Goode, D., Elgar, G., Hill, R.E., and Graaff, E. de (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Human Molecular Genetics *12*, 1725–1735.

Levings, P.P., and Bungert, J. (2002). The human β-globin locus control region: A center of attraction. Eur J Biochem *269*, 1589–1599.

Li, G., Ruan, X., Auerbach, R.K., Sandhu, K.S., Zheng, M., Wang, P., Poh, H.M., Goh, Y., Lim, J., Zhang, J., et al. (2012). Extensive Promoter-Centered Chromatin Interactions Provide a Topological Basis for Transcription Regulation. Cell *148*, 84–98.

Li, P., Cavallero, S., Gu, Y., Chen, T.H.P., Hughes, J., Hassan, A.B., Brüning, J.C., Pashmforoush, M., and Sucov, H.M. (2011). IGF signaling directs ventricular cardiomyocyte proliferation during embryonic heart development. Dev Camb Engl *138*, 1795–1805.

Li, W., Notani, D., Ma, Q., Tanasa, B., Nunez, E., Chen, A.Y., Merkurjev, D., Zhang, J., Ohgi, K., Song, X., et al. (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature *498*, 516–520.

Li, W., Notani, D., and Rosenfeld, M.G. (2016). Enhancers as non-coding RNA transcription units: recent insights and future perspectives. Nature Reviews Genetics *17*, 207.

Liang, J., Zhou, H., Gerdt, C., Tan, M., Colson, T., Kaye, K.M., Kieff, E., and Zhao, B. (2016). Epstein–Barr virus super-enhancer eRNAs are essential for MYC oncogene expression and lymphoblast proliferation. Proc National Acad Sci *113*, 14121–14126.

Liang, X., Wang, G., Lin, L., Lowe, J., Zhang, Q., Bu, L., Chen, Y., Chen, J., Sun, Y., and Evans, S.M. (2013). HCN4 dynamically marks the first heart field and conduction system precursors. Circ Res *113*, 399–407.

Lieberman-Aiden, E., Berkum, N.L. van, Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science *326*, 289–293.

Lie-Venema, H., Akker, N.M.S. van den, Bax, N.A.M., Winter, E.M., Maas, S., Kekarainen, T., Hoeben, R.C., deRuiter, M.C., Poelmann, R.E., and Groot, A.C.G. (2007). Origin, Fate, and Function of Epicardium-Derived Cells (EPDCs) in Normal and Abnormal Cardiac Development. Sci World J *7*, 1777–1798.

Lim, L.W.K., Chung, H.H., Chong, Y.L., and Lee, N.K. (2018). A survey of recently emerged genome-wide computational enhancer predictor tools. Comput Biol Chem *74*, 132–141.

Lints, T.J., Parsons, L.M., Hartley, L., Lyons, I., and Harvey, R.P. (1993). Nkx-2.5: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants. Development *119*, 419–431.

Liu, Q., Jiang, C., Xu, J., Zhao, M.-T., Bortle, K.V., Cheng, X., Wang, G., Chang, H.Y., Wu, J.C., and Snyder, M.P. (2017a). Genome-Wide Temporal Profiling of Transcriptome and Open Chromatin of Early Cardiomyocyte Differentiation Derived From hiPSCs and hESCs. Circ Res *121*, 376–391.

Liu, R., Lee, J., Kim, B.S., Wang, Q., Buxton, S.K., Balasubramanyam, N., Kim, J.J., Dong, J., Zhang, A., Li, S., et al. (2017b). Tead1 is required for maintaining adult cardiomyocyte function, and its loss results in lethal dilated cardiomyopathy. Jci Insight *2*, e93343.

Liu, Y., Asakura, M., Inoue, H., Nakamura, T., Sano, M., Niu, Z., Chen, M., Schwartz, R.J., and Schneider, M.D. (2007). Sox17 is essential for the specification of cardiac mesoderm in embryonic stem cells. Proc National Acad Sci *104*, 3859–3864.

Local, A., Huang, H., Albuquerque, C.P., Singh, N., Lee, A.Y., Wang, W., Wang, C., Hsia, J.E., Shiau, A.K., Ge, K., et al. (2017). Identification of H3K4me1-associated proteins at mammalian enhancers. Nat Genet *50*, 73–82.

Long, H.K., Prescott, S.L., and Wysocka, J. (2016). Ever-Changing Landscapes: Transcriptional

Enhancers in Development and Evolution. Cell *167*, 1170–1187.

Lovén, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers. Cell *153*, 320–334.

Luna-Zurita, L., Stirnimann, C.U., Glatt, S., Kaynak, B.L., Thomas, S., Baudin, F., Samee, M.A.H., He, D., Small, E.M., Mileikovsky, M., et al. (2016). Complex Interdependence Regulates Heterotypic Transcription Factor Distribution and Coordinates Cardiogenesis. Cell *164*, 999–1014.

Lyons, I., Parsons, L.M., Hartley, L., Li, R., Andrews, J.E., Robb, L., and Harvey, R.P. (1995). Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. Genes & Development *9*, 1654–1666.

Macleod, D., Charlton, J., Mullins, J., and Bird, A.P. (1994). Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. Genes & Development *8*, 2282–2292.

Maeda, T., Gupta, M.P., and Stewart, A.F.R. (2002a). TEF-1 and MEF2 transcription factors interact to regulate muscle-specific promoters. Biochem Bioph Res Co *294*, 791–797.

Maeda, T., Mazzulli, J.R., Farrance, I.K.G., and Stewart, A.F.R. (2002b). Mouse DTEF-1 (ETFR-1, TEF-5) Is a Transcriptional Activator in α 1 -Adrenergic Agonist-stimulated Cardiac Myocytes. J Biol Chem *277*, 24346–24352.

Makita, R., Uchijima, Y., Nishiyama, K., Amano, T., Chen, Q., Takeuchi, T., Mitani, A., Nagase, T., Yatomi, Y., Aburatani, H., et al. (2008). Multiple renal cysts, urinary concentration defects, and pulmonary emphysematous changes in mice lacking TAZ. Am J Physiology Ren Physiology *294*, F542-53.

Mansour, M.R., Abraham, B.J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A.D., Etchin, J., Lawton, L., Sallan, S.E., Silverman, L.B., et al. (2014). An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. Science *346*, 1373–1377.

Maston, G.A., Evans, S.K., and Green, M.R. (2006). Transcriptional Regulatory Elements in the Human Genome. Annual Review of Genomics and Human Genetics *7*, 29–59.

Mattioli, K., Volders, P.-J., Gerhardinger, C., Lee, J.C., Maass, P.G., Melé, M., and Rinn, J.L. (2019). High-throughput functional analysis of lncRNA core promoters elucidates rules governing tissue specificity. Genome Res *29*, 344–355.

Maurano, M.T., Wang, H., John, S., Shafer, A., Canfield, T., Lee, K., and Stamatoyannopoulos, J.A. (2015). Role of DNA Methylation in Modulating Transcription Factor Occupancy. Cell

Reports *12*, 1184–1195.

May, D., Blow, M.J., Kaplan, T., McCulley, D.J., Jensen, B.C., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., et al. (2011). Large-scale discovery of enhancers from human heart tissue. Nature Genetics *44*, 89–93.

Mayran, A., and Drouin, J. (2018). Pioneer transcription factors shape the epigenetic landscape. J Biol Chem *293*, 13795–13804.

Mayran, A., Sochodolsky, K., Khetchoumian, K., Harris, J., Gauthier, Y., Bemmo, A., Balsalobre, A., and Drouin, J. (2019). Pioneer and nonpioneer factor cooperation drives lineage specific chromatin opening. Nat Commun *10*, 3807.

McFadden, D.G., and Olson, E.N. (2002). Heart development: learning from mistakes. Current Opinion in Genetics & Development *12*, 328–335.

McGinty, R.K., and Tan, S. (2014). Nucleosome structure and function. Chem Rev *115*, 2255–2273.

McKinnell, I.W., Ishibashi, J., Grand, F.L., Punch, V.G.J., Addicks, G.C., Greenblatt, J.F., Dilworth, F.J., and Rudnicki, M.A. (2007). Pax7 activates myogenic genes by recruitment of a histone methyltransferase complex. Nat Cell Biol *10*, 77–84.

Meilhac, S.M., and Buckingham, M.E. (2018). The deployment of cell lineages that form the mammalian heart. Nat Rev Cardiol *15*, 705–724.

Meilhac, S.M., Esner, M., Kelly, R.G., Nicolas, J.-F., and Buckingham, M.E. (2004). The Clonal Origin of Myocardial Cells in Different Regions of the Embryonic Mouse Heart. Developmental Cell *6*, 685–698.

Melnikov, A., Murugan, A., Zhang, X., Tesileanu, T., Wang, L., Rogov, P., Feizi, S., Gnirke, A., Callan, C.G., Kinney, J.B., et al. (2012). Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. Nature Biotechnology *30*, 271–277.

Melo, C.A., Drost, J., Wijchers, P.J., Werken, H.V.D., Wit, E. de, Vrielink, J. a F.O., Elkon, R., Melo, S.A., Léveillé, N., Kalluri, R., et al. (2013). eRNAs Are Required for p53-Dependent Enhancer Activity and Gene Transcription. Molecular Cell *49*, 524–535.

Mendenhall, E.M., Williamson, K.E., Reyon, D., Zou, J.Y., Ram, O., Joung, J.K., and Bernstein, B.E. (2013). Locus-specific editing of histone modifications at endogenous enhancers. Nature Biotechnology *31*, 1133–1136.

Merika, M., Williams, A.J., Chen, G., Collins, T., and Thanos, D. (1998). Recruitment of CBP/p300

by the IFNβ Enhanceosome Is Required for Synergistic Activation of Transcription. Mol Cell *1*, 277–287.

Miesfeld, J.B., and Link, B.A. (2014). Establishment of transgenic lines to monitor and manipulate Yap/Taz-Tead activity in zebrafish reveals both evolutionarily conserved and divergent functions of the Hippo pathway. Mechanisms of Development *133*, 177–188.

Mifsud, B., Tavares-Cadete, F., Young, A.N., Sugar, R., Schoenfelder, S., Ferreira, L., Wingett, S.W., Andrews, S., Grey, W., Ewels, P.A., et al. (2015). Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. Nature Genetics *47*, 598–606.

Mikawa, T., and Hurtado, R. (2007). Development of the cardiac conduction system. Semin Cell Dev Biol *18*, 90–100.

Milewski, R.C., Chi, N.C., Li, J., Brown, C., Lu, M.M., and Epstein, J.A. (2004). Identification of minimal enhancer elements sufficient for Pax3 expression in neural crest and implication of Tead2 as a regulator of Pax3. Development *131*, 829–837.

Min, I.M., Waterfall, J.J., Core, L.J., Munroe, R.J., Schimenti, J., and Lis, J.T. (2011). Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. Gene Dev *25*, 742–754.

Mirtschink, P., Bischof, C., Pham, M.-D., Sharma, R., Khadayate, S., Rossi, G., Fankhauser, N., Traub, S., Sossalla, S., Hagag, E., et al. (2019). Inhibition of the Hypoxia-Inducible Factor 1α–Induced Cardiospecific HERNA1 Enhance-Templated RNA Protects From Heart Disease. Circulation *139*, 2778–2792.

Misfeldt, A.M., Boyle, S.C., Tompkins, K.L., Bautch, V.L., Labosky, P.A., and Baldwin, H.S. (2009). Endocardial cells are a distinct endothelial lineage derived from Flk1+ multipotent cardiovascular progenitors. Developmental Biology *333*, 78–89.

Mivelaz, M., Cao, A.-M., Kubik, S., Zencir, S., Hovius, R., Boichenko, I., Stachowicz, A.M., Kurat, C.F., Shore, D., and Fierz, B. (2019). Chromatin Fiber Invasion and Nucleosome Displacement by the Rap1 Transcription Factor. Mol Cell *77*, 488-500.e9.

Mo, J.-S., Park, H.W., and Guan, K.-L. (2014). The Hippo signaling pathway in stem cell biology and cancer. EMBO Reports *15*, 642–656.

Mohan, R.A., Boukens, B.J., and Christoffels, V.M. (2018). Developmental Origin of the Cardiac Conduction System: Insight from Lineage Tracing. Pediatr Cardiol *39*, 1107–1114.

Montefiori, L.E., Sobreira, D.R., Sakabe, N.J., Aneas, I., Joslin, A.C., Hansen, G.T., Bozek, G., Moskowitz, I.P., McNally, E.M., and Nóbrega, M.A. (2018). A promoter interaction map for cardiovascular disease genetics. Elife *7*, e35788.

Moreau, P., Hen, R., Wasylyk, B., Everett, R., Gaub, M.P., and Chambon, P. (1981). The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. Nucleic Acids Research *9*, 6047–6068.

Moretti, A., Caron, L., Nakano, A., Lam, J.T., Bernshausen, A., Chen, Y., Qyang, Y., Bu, L., Sasaki, M., Martin-Puig, S., et al. (2006). Multipotent Embryonic Isl1+ Progenitor Cells Lead to Cardiac, Smooth Muscle, and Endothelial Cell Diversification. Cell *127*, 1151–1165.

Mosqueira, D., Pagliari, S., Uto, K., Ebara, M., Romanazzo, S., Escobedo-Lucea, C., Nakanishi, J., Taniguchi, A., Franzese, O., Nardo, P.D., et al. (2014). Hippo Pathway Effectors Control Cardiac Progenitor Cell Fate by Acting as Dynamic Sensors of Substrate Mechanics and Nanostructure. ACS Nano *8*, 2033–2047.

Mumbach, M.R., Satpathy, A.T., Boyle, E.A., Dai, C., Gowen, B.G., Cho, S.W., Nguyen, M.L., Rubin, A.J., Granja, J.M., Kazane, K.R., et al. (2017). Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. Nat Genet *49*, 1602–1612.

Murakami, M., Nakagawa, M., Olson, E.N., and Nakagawa, O. (2005). A WW domain protein TAZ is a critical coactivator for TBX5, a transcription factor implicated in Holt-Oram syndrome. Proceedings of the National Academy of Sciences *102*, 18034–18039.

Nair, S.J., Yang, L., Meluzzi, D., Oh, S., Yang, F., Friedman, M.J., Wang, S., Suter, T., Alshareedah, I., Gamliel, A., et al. (2019). Phase separation of ligand-activated enhancers licenses cooperative chromosomal enhancer assembly. Nat Struct Mol Biol *26*, 193–203.

Naito, A.T., Shiojima, I., Akazawa, H., Hidaka, K., Morisaki, T., Kikuchi, A., and Komuro, I. (2006). Developmental stage-specific biphasic roles of Wnt/beta-catenin signaling in cardiomyogenesis and hematopoiesis. Proc National Acad Sci *103*, 19812–19817.

Nakano, A., Nakano, H., Smith, K.A., and Palpant, N.J. (2016). The developmental origins and lineage contributions of endocardial endothelium. Biochimica Et Biophysica Acta Bba - Mol Cell Res *1863*, 1937–1947.

Narlikar, L., Sakabe, N.J., Blanski, A.A., Arimura, F.E., Westlund, J.M., Nobrega, M.A., and Ovcharenko, I. (2010). Genome-wide discovery of human heart enhancers. Genome Research *20*, 381–392.

Naville, M., Ishibashi, M., Ferg, M., Bengani, H., Rinkwitz, S., Krecsmarik, M., Hawkins, T.A., Wilson, S.W., Manning, E., Chilamakuri, C.S.R., et al. (2015). Long-range evolutionary constraints reveal cis-regulatory interactions on the human X chromosome. Nat Commun *6*, 6904.

Nishioka, N., Yamamoto, S., Kiyonari, H., Sato, H., Sawada, A., Ota, M., Nakao, K., and Sasaki, H. (2008). Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. Mechanisms of Development *125*, 270–283.

Noseda, M., Peterkin, T., Simões, F.C., Patient, R., and Schneider, M.D. (2011). Cardiopoietic Factors. Circ Res *108*, 129–152.

Ohno, S. (1979). Monographs on Endocrinology.

O'Kane, C.J., and Gehring, W.J. (1987). Detection in situ of genomic regulatory elements in Drosophila. Proc National Acad Sci *84*, 9123–9127.

Olson, E.N. (2006). Gene Regulatory Networks in the Evolution and Development of the Heart. Science *313*, 1922–1927.

O'Meara, C.C., Wamstad, J.A., Gladstone, R.A., Fomovsky, G.M., Butty, V.L., Shrikumar, A., Gannon, J.B., Boyer, L.A., and Lee, R.T. (2015). Transcriptional reversion of cardiac myocyte fate during mammalian cardiac regeneration. Circulation Research *116*, 804–815.

Ong, C.-T., and Corces, V.G. (2014). CTCF: an architectural protein bridging genome topology and function. Nat Rev Genet *15*, 234–246.

Orom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytnicki, M., Notredame, C., Huang, Q., et al. (2010). Long Noncoding RNAs with Enhancer-like Function in Human Cells. Cell *143*, 46–58.

Ota, M., and Sasaki, H. (2008). Mammalian Tead proteins regulate cell proliferation and contact inhibition as transcriptional mediators of Hippo signaling. Development *135*, 4059–4069.

Øyen, N., Poulsen, G., Boyd, H.A., Wohlfahrt, J., Jensen, P.K.A., and Melbye, M. (2009). Recurrence of congenital heart defects in families. Circulation *120*, 295–301.

Pan, D. (2007). Hippo signaling in organ size control. Genes & Development *21*, 886–897.

Paramasivam, M., Sarkeshik, A., Yates, J.R., Fernandes, M.J.G., and McCollum, D. (2011). Angiomotin family proteins are novel activators of the LATS2 kinase tumor suppressor. Molecular Biology of the Cell *22*, 3725–3733.

Parker, S.C.J., Stitzel, M.L., Taylor, D.L., Orozco, J.M., Erdos, M.R., Akiyama, J.A., Bueren, K.L. van, Chines, P.S., Narisu, N., Program, N.C.S., et al. (2013). Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. Proceedings of the National Academy of Sciences *110*, 17921–17926.

Patwardhan, R.P., Hiatt, J.B., Witten, D.M., Kim, M.J., Smith, R.P., May, D., Lee, C., Andrie, J.M., Lee, S.-I., Cooper, G.M., et al. (2012). Massively parallel functional dissection of mammalian enhancers in vivo. Nature Biotechnology *30*, 265–270.

Petrovic, J., Zhou, Y., Fasolino, M., Goldman, N., Schwartz, G.W., Mumbach, M.R., Nguyen, S.C.,

Rome, K.S., Sela, Y., Zapataro, Z., et al. (2019). Oncogenic Notch Promotes Long-Range Regulatory Interactions within Hyperconnected 3D Cliques. Mol Cell *73*, 1174-1190.e12.

Pezone, A., Zuchegna, C., Tramontano, A., Romano, A., Russo, G., Rosa, M. de, Vinciguerra, M., Porcellini, A., Gottesman, M.E., and Avvedimento, E.V. (2019). RNA Stabilizes Transcription-Dependent Chromatin Loops Induced By Nuclear Hormones. Sci Rep-Uk *9*, 3925.

Pfeufer, A., Noord, C. van, Marciante, K.D., Arking, D.E., Larson, M.G., Smith, A.V., Tarasov, K.V., Müller, M., Sotoodehnia, N., Sinner, M.F., et al. (2010). Genome-wide association study of PR interval. Nature Genetics *42*, 153–159.

Phanstiel, D.H., Bortle, K.V., Spacek, D., Hess, G.T., Shamim, M.S., Machol, I., Love, M.I., Aiden, E.L., Bassik, M.C., and Snyder, M.P. (2017). Static and Dynamic DNA Loops form AP-1-Bound Activation Hubs during Macrophage Development. Mol Cell *67*, 1037-1048.e6.

Phillips-Cremins, J.E., Sauria, M.E.G., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S.K., Ong, C.-T., Hookway, T.A., Guo, C., Sun, Y., et al. (2013). Architectural Protein Subclasses Shape 3D Organization of Genomes during Lineage Commitment. Cell *153*, 1281–1295.

Pique-Regi, R., Degner, J.F., Pai, A.A., Gaffney, D.J., Gilad, Y., and Pritchard, J.K. (2010). Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data. Genome Res *21*, 447–455.

Pnueli, L., Rudnizky, S., Yosefzon, Y., and Melamed, P. (2015). RNA transcribed from a distal enhancer is required for activating the chromatin at the promoter of the gonadotropin α-subunit gene. Proc National Acad Sci *112*, 4369–4374.

Pobbati, A.V., and Hong, W. (2014). Emerging roles of TEAD transcription factors and its coactivators in cancers. Cancer Biology & Therapy *14*, 390–398.

Postma, A.V., Bezzina, C.R., and Christoffels, V.M. (2016). Genetics of congenital heart disease: the contribution of the noncoding regulatory genome. Journal of Human Genetics *61*, 13–19.

Prendiville, T., Jay, P.Y., and Pu, W.T. (2014). Insights into the Genetic Structure of Congenital Heart Disease from Human and Murine Studies on Monogenic Disorders. Cold Spring Harbor Perspectives in Medicine *4*, a013946–a013946.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell *152*, 1173–1183.

Quenneville, S., Verde, G., Corsinotti, A., Kapopoulou, A., Jakobsson, J., Offner, S., Baglivo, I., Pedone, P.V., Grimaldi, G., Riccio, A., et al. (2011). In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting

control regions. Molecular Cell *44*, 361–372.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2010). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279–283.

Rada-Iglesias, A., Bajpai, R., Prescott, S., Brugmann, S.A., Swigut, T., and Wysocka, J. (2012). Epigenomic Annotation of Enhancers Predicts Transcriptional Regulators of Human Neural Crest. Cell Stem Cell *11*, 633–648.

Ragni, C.V., Diguet, N., Garrec, J.-F.L., Novotova, M., Resende, T.P., Pop, S., Charon, N., Guillemot, L., Kitasato, L., Badouel, C., et al. (2017). Amotl1 mediates sequestration of the Hippo effector Yap1 downstream of Fat4 to restrict heart growth. Nat Commun *8*, 14582.

Raisner, R., Kharbanda, S., Jin, L., Jeng, E., Chan, E., Merchant, M., Haverty, P.M., Bainer, R., Cheung, T., Arnott, D., et al. (2018). Enhancer Activity Requires CBP/P300 Bromodomain-Dependent Histone H3K27 Acetylation. Cell Reports *24*, 1722–1729.

Rajagopal, N., Srinivasan, S., Kooshesh, K., Guo, Y., Edwards, M.D., Banerjee, B., Syed, T., Emons, B.J.M., Gifford, D.K., and Sherwood, R.I. (2016). High-throughput mapping of regulatory DNA. Nature Biotechnology *34*, 167–174.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell *159*, 1665–1680.

Rao, S.S.P., Huang, S.-C., Hilaire, B.G.S., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.-R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D., et al. (2017). Cohesin Loss Eliminates All Loop Domains. Cell *171*, 305-320.e24.

Reiter, F., Wienerroither, S., and Stark, A. (2017). Combinatorial function of transcription factors and cofactors. Curr Opin Genet Dev *43*, 73–81.

Ren, F., Zhang, L., and Jiang, J. (2009). Hippo signaling regulates Yorkie nuclear localization and activity through 14-3-3 dependent and independent mechanisms. Dev Biol *337*, 303–312.

Richards, A.A., and Garg, V. (2010). Genetics of Congenital Heart Disease. Current Cardiology Reviews *6*, 91–97.

Rickels, R., and Shilatifard, A. (2018). Enhancer Logic and Mechanics in Development and Disease. Trends Cell Biol *28*, 608–630.

Rishi, V., Bhattacharya, P., Chatterjee, R., Rozenberg, J., Zhao, J., Glass, K., Fitzgerald, P., and Vinson, C. (2010). CpG methylation of half-CRE sequences creates C/EBPα binding sites that

activate some tissue-specific genes. Proceedings of the National Academy of Sciences *107*, 20311–20316.

Roy, S., Siahpirani, A.F., Chasman, D., Knaack, S., Ay, F., Stewart, R., Wilson, M., and Sridharan, R. (2015). A predictive modeling approach for cell line-specific long-range regulatory interactions. Nucleic Acids Research *43*, 8694–8712.

Rubin, A.J., Barajas, B.C., Furlan-Magaril, M., Lopez-Pajares, V., Mumbach, M.R., Howard, I., Kim, D.S., Boxer, L.D., Cairns, J., Spivakov, M., et al. (2017). Lineage-specific dynamic and pre-established enhancer-promoter contacts cooperate in terminal differentiation. Nat Genet *49*, 1522–1528.

Sabari, B.R., Dall'Agnese, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K., Abraham, B.J., Hannett, N.M., Zamudio, A.V., Manteiga, J.C., et al. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. Science *361*, eaar3958.

Saga, Y., Miyagawa-Tomita, S., Takagi, A., Kitajima, S., Miyazaki, J. i, and Inoue, T. (1999). MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube. Dev Camb Engl *126*, 3437–3447.

Sahlén, P., Abdullayev, I., Ramsköld, D., Matskova, L., Rilakovic, N., Lötstedt, B., Albert, T.J., Lundeberg, J., and Sandberg, R. (2015). Genome-wide mapping of promoter-anchored interactions with close to single-enhancer resolution. Genome Biol *16*, 156.

Saint-André, V., Federation, A.J., Lin, C.Y., Abraham, B.J., Reddy, J., Lee, T.I., Bradner, J.E., and Young, R.A. (2016). Models of human core transcriptional regulatory circuitries. Genome Res *26*, 385–396.

Sanjana, N.E., Wright, J., Zheng, K., Shalem, O., Fontanillas, P., Joung, J., Cheng, C., Regev, A., and Zhang, F. (2016). High-resolution interrogation of functional elements in the noncoding genome. Science *353*, 1545–1549.

Sardina, J.L., Collombet, S., Tian, T.V., Gómez, A., Stefano, B.D., Berenguer, C., Brumbaugh, J., Stadhouders, R., Segura-Morales, C., Gut, M., et al. (2018). Transcription Factors Drive Tet2-Mediated Enhancer Demethylation to Reprogram Cell Fate. Cell Stem Cell *23*, 727-741.e9.

Sawada, A., Kiyonari, H., Ukita, K., Nishioka, N., Imuta, Y., and Sasaki, H. (2008). Redundant roles of Tead1 and Tead2 in notochord development and the regulation of cell proliferation and survival. Mol Cell Biol *28*, 3177–3189.

Schaukowitch, K., Joo, J.-Y., Liu, X., Watts, J.K., Martinez, C., and Kim, T.-K. (2014). Enhancer RNA facilitates NELF release from immediate early genes. Molecular Cell *56*, 29–42.

Schmitt, A.D., Hu, M., and Ren, B. (2016). Genome-wide mapping and analysis of chromosome

architecture. Nat Rev Mol Cell Bio *17*, 743–755.

Schneider, T.D., and Stephens, R.M. (1990). Sequence logos: a new way to display consensus sequences. Nucleic Acids Res *18*, 6097–6100.

Senyo, S.E., Steinhauser, M.L., Pizzimenti, C.L., Yang, V.K., Cai, L., Wang, M., Wu, T.-D., Guerquin-Kern, J.-L., Lechene, C.P., and Lee, R.T. (2013). Mammalian heart renewal by pre-existing cardiomyocytes. Nature *493*, 433–436.

Serandour, A.A., Avner, S., Percevault, F., Demay, F., Bizot, M., Lucchetti-Miganeh, C., Barloy-Hubler, F., Brown, M., Lupien, M., Metivier, R., et al. (2011). Epigenetic switch involved in activation of pioneer factor FOXA1-dependent enhancers. Genome Res *21*, 555–565.

Seyres, D., Ghavi-Helm, Y., Junion, G., Taghli-Lamallem, O., Guichard, C., Röder, L., Girardot, C., Furlong, E.E.M., and Perrin, L. (2016). Identification and in silico modeling of enhancers reveals new features of the cardiac differentiation network. Development *143*, 4533–4542.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., et al. (2012). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116–120.

Sherwood, R.I., Hashimoto, T., O'Donnell, C.W., Lewis, S., Barkal, A.A., Hoff, J.P. van, Karun, V., Jaakkola, T., and Gifford, D.K. (2014). Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. Nature Biotechnology *32*, 171–178.

Shii, L., Song, L., Maurer, K., Zhang, Z., and Sullivan, K.E. (2017). SERPINB2 is regulated by dynamic interactions with pause-release proteins and enhancer RNAs. Mol Immunol *88*, 20–31.

Shin, H.Y., Willi, M., Yoo, K.H., Zeng, X., Wang, C., Metser, G., and Hennighausen, L. (2016). Hierarchy within the mammary STAT5-driven Wap super-enhancer. Nat Genet *48*, 904–911.

Shuaib, M., Parsi, K.M., Thimma, M., Adroub, S.A., Kawaji, H., Seridi, L., Ghosheh, Y., Fort, A., Fallatah, B., Ravasi, T., et al. (2019). Nuclear AGO1 Regulates Gene Expression by Affecting Chromatin Architecture in Human Cells. Cell Syst.

Sifrim, A., Hitz, M.-P., Wilsdon, A., Breckpot, J., Turki, S.H.A., Thienpont, B., McRae, J., Fitzgerald, T.W., Singh, T., Swaminathan, G.J., et al. (2016). Distinct genetic architectures for syndromic and nonsyndromic congenital heart defects identified by exome sequencing. Nat Genet *48*, 1060–1065.

Sigova, A.A., Abraham, B.J., Ji, X., Molinie, B., Hannett, N.M., Guo, Y.E., Jangi, M., Giallourakis, C.C., Sharp, P. a, and Young, R.A. (2015). Transcription factor trapping by RNA in gene regulatory elements. Science *350*, 978–981.

Simeonov, D.R., Gowen, B.G., Boontanrart, M., Roth, T.L., Gagnon, J.D., Mumbach, M.R., Satpathy, A.T., Lee, Y., Bray, N.L., Chan, A.Y., et al. (2017). Discovery of stimulation-responsive immune enhancers with CRISPR activation. Nature *549*, 111–115.

Simpson, P. (1983). Norepinephrine-stimulated hypertrophy of cultured rat myocardial cells is an alpha 1 adrenergic response. Journal of Clinical Investigation *72*, 732–738.

Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., Zhou, T., Rohs, R., Honig, B., Bussemaker, H.J., et al. (2011). Cofactor Binding Evokes Latent Differences in DNA Binding Specificity between Hox Proteins. Cell *147*, 1270–1282.

Slattery, M., Zhou, T., Yang, L., Machado, A.C.D., Gordân, R., and Rohs, R. (2014). Absence of a simple code: how transcription factors read the genome. Trends Biochem Sci *39*, 381–399.

Smemo, S., Campos, L.C., Moskowitz, I.P., Krieger, J.E., Pereira, A.C., and Nobrega, M.A. (2012). Regulatory variation in a TBX5 enhancer leads to isolated congenital heart disease. Human Molecular Genetics *21*, 3255–3263.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nat Rev Genetics *14*, 204–220.

Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell *151*, 994–1004.

Soufi, A., Garcia, M.F., Jaroszewicz, A., Osman, N., Pellegrini, M., and Zaret, K.S. (2015). Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. Cell *161*, 555–568.

Soysa, T.Y. de, Ranade, S.S., Okawa, S., Ravichandran, S., Huang, Y., Salunga, H.T., Schricker, A., Sol, A. del, Gifford, C.A., and Srivastava, D. (2019). Single-cell analysis of cardiogenesis reveals basis for organ-level developmental defects. Nature *572*, 120–124.

Später, D., Abramczuk, M.K., Buac, K., Zangi, L., Stachel, M.W., Clarke, J., Sahara, M., Ludwig, A., and Chien, K.R. (2013). A HCN4+ cardiomyogenic progenitor derived from the first heart field and human pluripotent stem cells. Nat Cell Biol *15*, 1098–1106.

Später, D., Hansson, E.M., Zangi, L., and Chien, K.R. (2014). How to make a cardiomyocyte. Dev Camb Engl *141*, 4418–4431.

Spitz, F., and Furlong, E.E.M. (2012). Transcription factors: from enhancer binding to developmental control. Nature Reviews Genetics *13*, 613–626.

Srivastava, D. (2006). Making or Breaking the Heart: From Lineage Determination to Morphogenesis. Cell *126*, 1037–1048.

Staby, L., O'Shea, C., Willemoës, M., Theisen, F., Kragelund, B.B., and Skriver, K. (2017). Eukaryotic transcription factors: paradigms of protein intrinsic disorder. Biochem J *474*, 2509–2532.

Stergachis, A.B., Neph, S., Reynolds, A., Humbert, R., Miller, B., Paige, S.L., Vernot, B., Cheng, J.B., Thurman, R.E., Sandstrom, R., et al. (2013). Developmental Fate and Cellular Maturity Encoded in Human Regulatory DNA Landscapes. Cell *154*, 888–903.

Stevens, T.J., Lando, D., Basu, S., Atkinson, L.P., Cao, Y., Lee, S.F., Leeb, M., Wohlfahrt, K.J., Boucher, W., O'Shaughnessy-Kirwan, A., et al. (2017). 3D structures of individual mammalian genomes studied by single-cell Hi-C. Nature *544*, 59–64.

Stewart, A.F.R., Suzow, J., Kubota, T., Ueyama, T., and Chen, H.-H. (1998). Transcription Factor RTEF-1 Mediates α1-Adrenergic Reactivation of the Fetal Gene Program in Cardiac Myocytes. Circ Res *83*, 43–49.

Stormo, G.D., and Zhao, Y. (2010). Determining the specificity of protein-DNA interactions. Nat Rev Genetics *11*, 751–760.

Subramanian, V., Fields, P.A., and Boyer, L.A. (2015). H2A.Z: a molecular rheostat for transcriptional control. F1000prime Reports *7*, 01.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell *126*, 663–676.

Takaku, M., Grimm, S.A., Shimbo, T., Perera, L., Menafra, R., Stunnenberg, H.G., Archer, T.K., Machida, S., Kurumizaka, H., and Wade, P.A. (2016). GATA3-dependent cellular reprogramming requires activation-domain dependent recruitment of a chromatin remodeler. Genome Biol *17*, 36.

Takeuchi, J.K., and Bruneau, B.G. (2009). Directed transdifferentiation of mouse mesoderm to heart tissue by defined factors. Nature *459*, 708–711.

Talbert, P.B., and Henikoff, S. (2016). Histone variants on the move: substrates for chromatin dynamics. Nat Rev Mol Cell Bio *18*, 115–126.
Tam, P.P., Parameswaran, M., Kinder, S.J., and Weinberger, R.P. (1997). The allocation of epiblast cells to the embryonic heart and other mesodermal lineages: the role of ingression and tissue movement during gastrulation. Development *124*, 1631–1642.

Tan, S.H., Leong, W.Z., Ngoc, P.C.T., Tan, T.K., Bertulfo, F.C., Lim, M.C., An, O., Li, Z., Yeoh, A.E.J., Fullwood, M.J., et al. (2019). The enhancer RNA ARIEL activates the oncogenic transcriptional program in T-cell acute lymphoblastic leukemia. Blood *134*, 239–251.

Tapon, N., Harvey, K.F., Bell, D.W., Wahrer, D.C.R., Schiripo, T.A., Haber, D.A., and Hariharan,

I.K. (2002). salvador Promotes both cell cycle exit and apoptosis in Drosophila and is mutated in human cancer cell lines. Cell *110*, 467–478.

Tapscott, S., Davis, R., Thayer, M., Cheng, P., Weintraub, H., and Lassar, A. (1988). MyoD1: a nuclear phosphoprotein requiring a Myc homology region to convert fibroblasts to myoblasts. Science *242*, 405–411.

Tayal, U., Prasad, S., and Cook, S.A. (2017). Genetics and genomics of dilated cardiomyopathy and systolic heart failure. Genome Med *9*, 20.

Thanos, D., and Maniatis, T. (1995). Virus induction of human IFNβ gene expression requires the assembly of an enhanceosome. Cell *83*, 1091–1100.

Thompson, M., Andrade, V.A., Andrade, S.J., Pusl, T., Ortega, J.M., Goes, A.M., and Leite, M.F. (2003). Inhibition of the TEF/TEAD transcription factor activity by nuclear calcium and distinct kinase pathways. Biochemical and Biophysical Research Communications *301*, 267–274.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. Nature *489*, 75–82.

Tie, F., Banerjee, R., Stratton, C.A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M.O., Scacheri, P.C., and Harte, P.J. (2009). CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. Development *136*, 3131–3141.

Ueno, S., Weidinger, G., Osugi, T., Kohn, A.D., Golob, J.L., Pabon, L., Reinecke, H., Moon, R.T., and Murry, C.E. (2007). Biphasic role for Wnt/beta-catenin signaling in cardiac specification in zebrafish and embryonic stem cells. Proceedings of the National Academy of Sciences *104*, 9685–9690.

Uosaki, H., Cahan, P., Lee, D.I., Wang, S., Miyamoto, M., Fernandez, L., Kass, D.A., and Kwon, C. (2015). Transcriptional Landscape of Cardiomyocyte Maturation. Cell Reports *13*, 1705–1716.

Vandel, J., Cassan, O., Lèbre, S., Lecellier, C.-H., and Bréhélin, L. (2019). Probing transcription factor combinatorics in different promoter classes and in enhancers. Bmc Genomics *20*, 103.

Vanzan, L., Soldati, H., Ythier, V., Anand, S., Francis, N.J., and Murr, R. (2020). High throughput screening identifies SOX2 as a Super Pioneer Factor that inhibits DNA methylation maintenance at its binding sites. Biorxiv 2020.02.10.941682.

Vierbuchen, T., Ling, E., Cowley, C.J., Couch, C.H., Wang, X., Harmin, D.A., Roberts, C.W.M., and Greenberg, M.E. (2017). AP-1 Transcription Factors and the BAF Complex Mediate Signal-Dependent Enhancer Selection. Molecular Cell *68*, 1067-1082.e12.

Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L.A. (2007). VISTA Enhancer Browser--a database of tissue-specific human enhancers. Nucleic Acids Research *35*, D88–D92.

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature *457*, 854–858.

Wamstad, J.A., Alexander, J.M., Truty, R.M., Shrikumar, A., Li, F., Eilertson, K.E., Ding, H., Wylie, J.N., Pico, A.R., Capra, J.A., et al. (2012). Dynamic and Coordinated Epigenetic Regulation of Developmental Transitions in the Cardiac Lineage. Cell *151*, 206–220.

Wamstad, J.A., Wang, X., Demuren, O.O., and Boyer, L.A. (2014). Distal enhancers: new insights into heart development and disease. Trends in Cell Biology *24*, 294–302.

Wang, H., Maurano, M.T., Qu, H., Varley, K.E., Gertz, J., Pauli, F., Lee, K., Canfield, T., Weaver, M., Sandstrom, R., et al. (2012). Widespread plasticity in CTCF occupancy linked to DNA methylation. Genome Research *22*, 1680–1688.

Wang, J., Liu, S., Heallen, T., and Martin, J.F. (2018). The Hippo pathway in the heart: pivotal roles in development, disease, and regeneration. Nat Rev Cardiol *15*, 672–684.

Wang, L., Du, Y., Ward, J.M., Shimbo, T., Lackford, B., Zheng, X., Miao, Y., Zhou, B., Han, L., Fargo, D.C., et al. (2014). INO80 Facilitates Pluripotency Gene Activation in Embryonic Stem Cell Self-Renewal, Reprogramming, and Blastocyst Development. Cell Stem Cell *14*, 575–591.

Wang, X., Tucker, N.R., Rizki, G., Mills, R., Krijger, P.H., Wit, E. de, Subramanian, V., Bartell, E., Nguyen, X.-X., Ye, J., et al. (2016). Discovery and validation of sub-threshold genome-wide association study loci using epigenomic signatures. ELife *5*, 1061.

Wang, X., Cairns, M.J., and Yan, J. (2019). Super-enhancers in transcriptional regulation and genome organization. Nucleic Acids Res *47*, 11481–11496.

Waspe, L.E., Ordahl, C.P., and Simpson, P.C. (1990). The cardiac beta-myosin heavy chain isogene is induced selectively in alpha 1-adrenergic receptor-stimulated hypertrophy of cultured rat heart myocytes. Journal of Clinical Investigation *85*, 1206–1214.

Weintraub, A.S., Li, C.H., Zamudio, A.V., Sigova, A.A., Hannett, N.M., Day, D.S., Abraham, B.J., Cohen, M.A., Nabet, B., Buckley, D.L., et al. (2017). YY1 Is a Structural Regulator of Enhancer-Promoter Loops. Cell *171*, 1573-1588.e28.

Weirauch, M.T., Cote, A., Norel, R., Annala, M., Zhao, Y., Riley, T.R., Saez-Rodriguez, J., Cokelaer, T., Vedenko, A., Talukder, S., et al. (2013). Evaluation of methods for modeling transcription factor sequence specificity. Nat Biotechnol *31*, 126–134.

Wessels, A., Hoff, M.J.B. van den, Adamo, R.F., Phelps, A.L., Lockhart, M.M., Sauls, K., Briggs, L.E., Norris, R.A., Wijk, B. van, Perez-Pomares, J.M., et al. (2012). Epicardially derived fibroblasts preferentially contribute to the parietal leaflets of the atrioventricular valves in the murine heart. Dev Biol *366*, 111–124.

Whalen, S., Truty, R.M., and Pollard, K.S. (2016). Enhancer–promoter interactions are encoded by complex genomic signatures on looping chromatin. Nature Genetics *48*, 488–496.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell *153*, 307–319.

Wilkinson, A.C., Nakauchi, H., and Göttgens, B. (2017). Mammalian Transcription Factor Networks: Recent Advances in Interrogating Biological Complexity. Cell Syst *5*, 319–331.

Wit, E. de, Bouwman, B.A.M., Zhu, Y., Klous, P., Splinter, E., Verstegen, M.J.A.M., Krijger, P.H.L., Festuccia, N., Nora, E.P., Welling, M., et al. (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. Nature *501*, 227–231.

Wu, C., Wong, Y.-C., and Elgin, S.C.R. (1979a). The chromatin structure of specific genes: II. Disruption of chromatin structure during gene activity. Cell *16*, 807–814.

Wu, C., Bingham, P.M., Livak, K.J., Holmgren, R., and Elgin, S.C.R. (1979b). The chromatin structure of specific genes: I. Evidence for higher order domains of defined DNA sequence. Cell *16*, 797–806.

Wu, S., Huang, J., Dong, J., and Pan, D. (2003). hippo encodes a Ste-20 family protein kinase that restricts cell proliferation and promotes apoptosis in conjunction with salvador and warts. Cell *114*, 445–456.

Xiao, J.H., Davidson, I., Ferrandon, D., Rosales, R., Vigneron, M., Macchi, M., Ruffenach, F., and Chambon, P. (1987). One cell-specific and three ubiquitous nuclear proteins bind in vitro to overlapping motifs in the domain B1 of the SV40 enhancer. The EMBO Journal *6*, 3005–3013.

Xiao, J.H., Davidson, I., Matthes, H., Garnier, J.-M., and Chambon, P. (1991). Cloning, expression, and transcriptional properties of the human enhancer factor TEF-1. Cell *65*, 551–568.

Xiao, Y., Hill, M.C., Zhang, M., Martin, T.J., Morikawa, Y., Wang, S., Moise, A.R., Wythe, J.D., and Martin, J.F. (2018). Hippo Signaling Plays an Essential Role in Cell State Transitions during Cardiac Fibroblast Development. Dev Cell *45*, 153-169.e6.

Xie, S., Duan, J., Li, B., Zhou, P., and Hon, G.C. (2017). Multiplexed Engineering and Analysis of Combinatorial Enhancer Activity in Single Cells. Mol Cell *66*, 285-299.e5.

Xin, M., Kim, Y., Sutherland, L.B., Qi, X., McAnally, J., Schwartz, R.J., Richardson, J.A., Bassel-Duby, R., and Olson, E.N. (2011). Regulation of Insulin-Like Growth Factor Signaling by Yap Governs Cardiomyocyte Proliferation and Embryonic Heart Size. Science Signaling *4*, ra70–ra70.

Yagi, R., Kohn, M.J., Karavanova, I., Kaneko, K.J., Vullhorst, D., DePamphilis, M.L., and Buonanno, A. (2007). Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. Development *134*, 3827–3836.

Yáñez-Cuna, J.O., Arnold, C.D., Stampfel, G., Boryń, L.M., Gerlach, D., Rath, M., and Stark, A. (2014). Dissection of thousands of cell type-specific enhancers identifies dinucleotide repeat motifs as general enhancer features. Genome Research *24*, 1147–1156.

Yao, L., Shen, H., Laird, P.W., Farnham, P.J., and Berman, B.P. (2015). Inferring regulatory element landscapes and transcription factor networks from cancer methylomes. Genome Biol *16*, 105.

Yu, F.-X., Zhao, B., and Guan, K.-L. (2015). Hippo Pathway in Organ Size Control, Tissue Homeostasis, and Cancer. Cell *163*, 811–828.

Yuan, X., Song, M., Devine, P., Bruneau, B.G., Scott, I.C., and Wilson, M.D. (2018). Heart enhancers with deeply conserved regulatory activity are established early in zebrafish development. Nat Commun *9*, 4977.

Zaidi, S., Choi, M., Wakimoto, H., Ma, L., Jiang, J., Overton, J.D., Romano-Adesman, A., Bjornson, R.D., Breitbart, R.E., Brown, K.K., et al. (2013). De novo mutations in histone-modifying genes in congenital heart disease. Nature *498*, 220–223.

Zeitlinger, J., Stark, A., Kellis, M., Hong, J.-W., Nechaev, S., Adelman, K., Levine, M., and Young, R.A. (2007). RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. Nature Genetics *39*, 1512–1516.

Zhang, H., Gan, H., Wang, Z., Lee, J.-H., Zhou, H., Ordog, T., Wold, M.S., Ljungman, M., and Zhang, Z. (2017). RPA Interacts with HIRA and Regulates H3.3 Deposition at Gene Regulatory Elements in Mammalian Cells. Mol Cell *65*, 272–284.
Zhang, H., Lui, K.O., and Zhou, B. (2018). Endocardial Cell Plasticity in Cardiac Development, Diseases and Regeneration. Circ Res *122*, 774–789.

Zhang, N., Bai, H., David, K.K., Dong, J., Zheng, Y., Cai, J., Giovannini, M., Liu, P., Anders, R.A., and Pan, D. (2010). The Merlin/NF2 tumor suppressor functions through the YAP oncoprotein to regulate tissue homeostasis in mammals. Developmental Cell *19*, 27–38.

Zhang, W., Prakash, C., Sum, C., Gong, Y., Li, Y., Kwok, J.J.T., Thiessen, N., Pettersson, S., Jones, S.J.M., Knapp, S., et al. (2012). Bromodomain-containing protein 4 (BRD4) regulates RNA polymerase II serine 2 phosphorylation in human CD4+ T cells. J Biological Chem *287*, 43137–

43155.

Zhang, Y., Zhang, D., Li, Q., Liang, J., Sun, L., Yi, X., Chen, Z., Yan, R., Xie, G., Li, W., et al. (2016). Nucleation of DNA repair factors by FOXA1 links DNA demethylation to transcriptional pioneering. Nat Genet *48*, 1003–1013.

Zhao, B., Wei, X., Li, W., Udan, R.S., Yang, Q., Kim, J., Xie, J., Ikenoue, T., Yu, J., Li, L., et al. (2007). Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control. Genes & Development *21*, 2747–2761.

Zhao, Y., Ruan, S., Pandey, M., and Stormo, G.D. (2012). Improved models for transcription factor binding site identification using nonindependent interactions. Genetics *191*, 781–790.

Zhao, Y., Wang, L., Ren, S., Wang, L., Blackburn, P.R., McNulty, M.S., Gao, X., Qiao, M., Vessella, R.L., Kohli, M., et al. (2016). Activation of P-TEFb by Androgen Receptor-Regulated Enhancer RNAs in Castration-Resistant Prostate Cancer. Cell Reports *15*, 599–610.

Zhao, Y., Zhou, J., He, L., Li, Y., Yuan, J., Sun, K., Chen, X., Bao, X., Esteban, M.A., Sun, H., et al. (2019). MyoD induced enhancer RNA interacts with hnRNPL to activate target gene transcription during myogenic differentiation. Nat Commun *10*, 5787.

Zhou, Q., Li, L., Zhao, B., and Guan, K.L. (2015a). The Hippo Pathway in Heart Development, Regeneration, and Diseases. Circulation Research *116*, 1431–1447.

Zhou, T., Shen, N., Yang, L., Abe, N., Horton, J., Mann, R.S., Bussemaker, H.J., Gordân, R., and Rohs, R. (2015b). Quantitative modeling of transcription factor binding specificities using DNA shape. Proc National Acad Sci *112*, 4654–4659.

Zhu, Y., Sun, L., Chen, Z., Whitaker, J.W., Wang, T., and Wang, W. (2013). Predicting enhancer transcription and activity from chromatin modifications. Nucleic Acids Res *41*, 10032–10043.

Zhu, Y., Chen, Z., Zhang, K., Wang, M., Medovoy, D., Whitaker, J.W., Ding, B., Li, N., Zheng, L., and Wang, W. (2016). Constructing 3D interaction maps from 1D epigenomes. Nat Commun *7*, 10812.

Chapter 2

TEAD1 regulates cardiac enhancer chromatin states

Author Contributions for Chapter 2:
All experiments were designed by Olukunle Demuren and Laurie Boyer, and performed by Olukunle Demuren, with the exception of the first ATAC-seq experiments, performed by Joseph Wamstad, and the ChIP-seq experiments, performed by Constantine Mylonas. The manuscript was prepared by Olukunle Demuren and Laurie Boyer.

Summary

Heart development requires precise control of gene programs as faulty regulation leads to congenital heart defects (CHD) and eventually heart failure, the leading cause of infant mortality and deaths worldwide, respectively. Although transcription factors (TFs) are known to control development by binding to regulatory elments such as enhancers, in most cases, we lack a fundamental understanding of how these factors coordinate regulate cell identity and function. To identify potential key regulatory TFs of heart development, we use the Assay for Transposase-Accessible Chromatin (ATAC-seq) to map changes in chromatin accessibility across the genome through several stages of embryonic stem cell (ESC) differentiation towards cardiomyocytes (CMs). We integrated these data with motif analysis, histone modification patterns, and gene expression to systematically identify driver TFs during cardiac lineage commitment. We found that TEA domain family member 1 (TEAD1) was highly enriched in CM-specific open chromatin in enhancers. Using an inducible degron construct, we show that TEAD1 loss led to abnormal beating phenotype, suggesting a key role in this process. Further mechanistic studies showed that TEAD1 is necessary to maintain open chromatin and active enhancer histone marks at a subset of enhancers that appear to regulate genes with roles in forming cell-cell contacts, distinct from gene expression programs regulated by other cardiac TFs. Together, these data demonstrate an integrated approach for identifying key regulators of lineage commitment and how TFs regulate distinct gene expression programs to ensure robust control of development.

INTRODUCTION

Transcription factors (TFs) translate signaling cues into defined transcriptional outcomes by binding to specific genomic sites and by recruiting regulatory complexes including chromatin modifiers and the transcriptional machinery (Purvis and Lahav, 2013; Reiter et al., 2017; Voss and Hager, 2013). Remarkably, although about 2,000 TFs are encoded in the mammalian genome, only a relatively small subset are thought to be critical regulators of tissue-specific gene expression and cell identity (Gerstein et al., 2012; Lambert et al., 2018; Neph et al., 2012; Sonawane et al., 2017; Spitz and Furlong, 2012). To date, lineage-specific TFs have been characterized in individual cell types, however, the dynamics of TF action during lineage commitment have been understudied. Thus, systematic and unbiased approaches to identify TF dynamics and novel linkages between TFs and their roles in developmental gene expression during cardiac lineage commitment are a critical next step for improving our understanding of development and the etiology of developmental disorders.

Heart development requires precise regulation of gene expression programs and faulty regulation leads to congenital heart defects (CHD) (Bruneau, 2013; Gelb and Chung, 2014). Although several cardiac TFs and chromatin regulators have been genetically linked to CHD (Homsy et al., 2015; Jin et al., 2017; Prendiville et al., 2014; Zaidi et al., 2013), how these factors control tissue-specific gene expression programs is not well-understood. In fact, most of the genetic causes of CHD remain unknown (Blue et al., 2017). Combined with the surprising complexity of the TF regulatory networks underpinning heart development (Akerberg et al., 2019; Luna-Zurita et al., 2016; Paige et al., 2012; Wamstad et al., 2012), these data suggest that many factors important for proper cardiac development have yet to be identified.

TFs often bind to *cis*-regulatory elements such as enhancers to activate the expression of specific genes by catalyzing the assembly of the transcriptional machinery at target promoters through looping mechanisms (Furlong and Levine, 2018). Enhancers have been identified genome-wide by the enrichment of specific chromatin marks such as the acetylation of lysine 27 of histone H3 (H3K27ac) and by the binding of clusters of lineage-specific TFs within non-

coding genomic regions including in heart (Akerberg et al., 2019; Creyghton et al., 2010; Gotea et al., 2010; Rada-Iglesias et al., 2012; Wamstad et al., 2012). Tens of thousands of enhancers have been characterized in individual cell types; however, knowledge of TFs that occupy these elements has been limited by candidate approaches such as ChIP-seq with antibodies to specific TFs or TFs tagged with epitope tags. Recent advances in mapping accessible chromatin using DNase- or ATAC-seq combined with bioinformatic methods provide an unbiased approach to identify candidate TFs that bind dynamically to the genome during developmental transitions (Buenrostro et al., 2013; Natarajan et al., 2012; Thurman et al., 2012). Importantly, genome-wide association studies (GWAS) show that common variants associated with specific cardiac traits can disrupt TF binding (Wang et al., 2016), such as in the case of the CHD Holt-Oram syndrome (Postma et al., 2016; Smemo et al., 2012). Thus, dissecting the TFs that regulate enhancer activity during cardiac fate transitions is crucial for developing a deeper understanding of the genetics of heart development and the processes that contribute to cardiac diseases.

Here, we measured chromatin accessibility genome-wide using ATAC-seq in a dynamic ESC-based model of cardiac lineage commitment (Wamstad et al., 2012). By combining these data with chromatin marks and transcriptional profiling, we find that the motif most closely matching the TEA domain (TEAD) family of TFs is enriched in a distinct subset of stage-specific enhancers in cardiac progenitors and cardiomyocytes. Further analysis suggested that TEAD1 is critical for cardiac lineage commitment based on its enrichment within super-enhancer clusters and its cardiac-specific expression. By generating an inducible degron-tagged version of TEAD1, we showed that its loss leads to an abnormal beating phenotype. Notably, TEAD1-depleted CMs showed decreased H3K27ac and a concomitant loss of chromatin accessibility at TEAD1-bound enhancers in proximity to genes involved in cell-cell contacts. Together, our study demonstrates an integrated approach for determining distinct roles of TFs during cardiac lineage commitment and identify TEAD1 as a potential regulator of intercellular junctions during heart development.

RESULTS

Dynamic chromatin accessibility marks enhancers in cardiac lineage commitment

To identify TFs in an unbiased manner that regulate cardiac-specific gene expression programs, we analyzed chromatin accessibility genome-wide during cardiomyocyte (CM) differentiation using the Assay for Transposase-Accessible Chromatin (ATAC-seq) protocol (Buenrostro et al., 2013, 2015). Chromatin and DNA accessibility often coincides with the binding of regulatory factors at discrete genomic loci (Buenrostro et al., 2013; Felsenfeld et al., 1996; Natarajan et al., 2012; Stergachis et al., 2013). Using an ESC-based differentiation strategy, we performed ATAC-seq during a continuous differentiation timecourse: ESCs (day 0), pre-cardiac mesoderm (MES; day 4), cardiac progenitors (CP; day 5.3) and CMs (day 10) (Wamstad et al., 2012) (**Figure 1A**). This differentiation protocol produces a highly enriched CM population, as well as defined intermediates representing progressive differentiation stages (Experimental Procedures). By deep sequencing of ATAC-seq libraries from multiple replicates at each stage of differentiation, we identified 171,542 unique open chromatin regions (**Figure 1B and Table 1**).

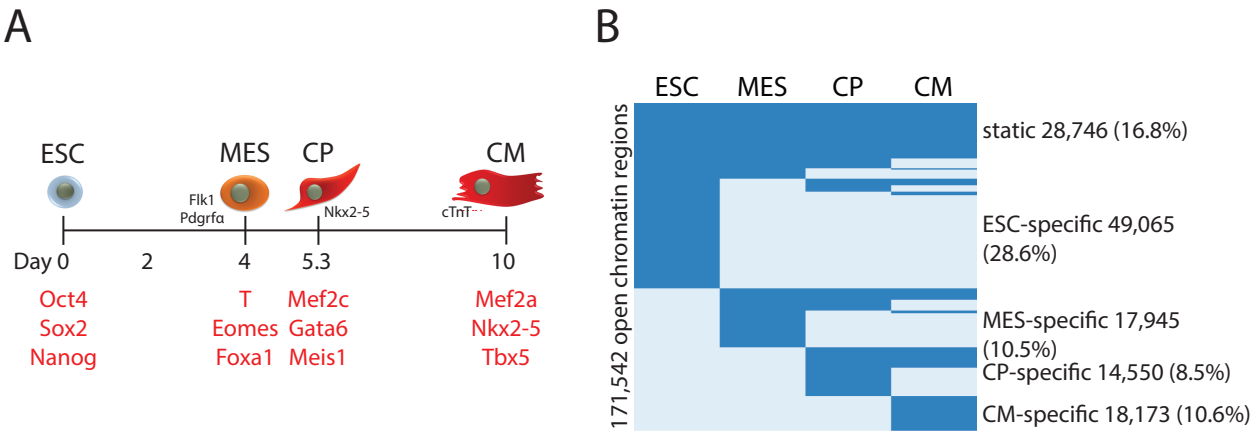


Figure 1: Dynamics of Chromatin Accessibility During Cardiomyocyte Differentiation

A) Schematic of *in vitro* cardiomyocyte (CM) differentiation. Embryonic stem cells (ESC) progress to pre-cardiac mesoderm (MES) in an embryoid body stage, and then differentiate to cardiac progenitors (CP) and cardiomyocytes (CM) in a monolayer upon treatment with the cytokines and growth factors indicated at top in blue. Key cell identity TFs expressed at each stage are indicated in red below the diagram. Factors used to assay the quality of differentiation are indicated in black.

B) Heatmap representing all open chromatin peaks across CM differentiation. Dark blue indicates an open peak in that stage, whereas light blue indicates a closed peak in that stage.

stage	total mapped reads (minus mitochondrial)	total peaks	stage-specific peaks	shared-specific peaks		static peaks
ESC	161,634,685	97,244	49,065	4,607		28,746
MES	151,161,322	69,983	17,945		5,285	
CP	173,800,731	77,546	14,550	10,835		
CM	196,356,140	71,598	18,173			

Table 1: Summary of ATAC-seq Data

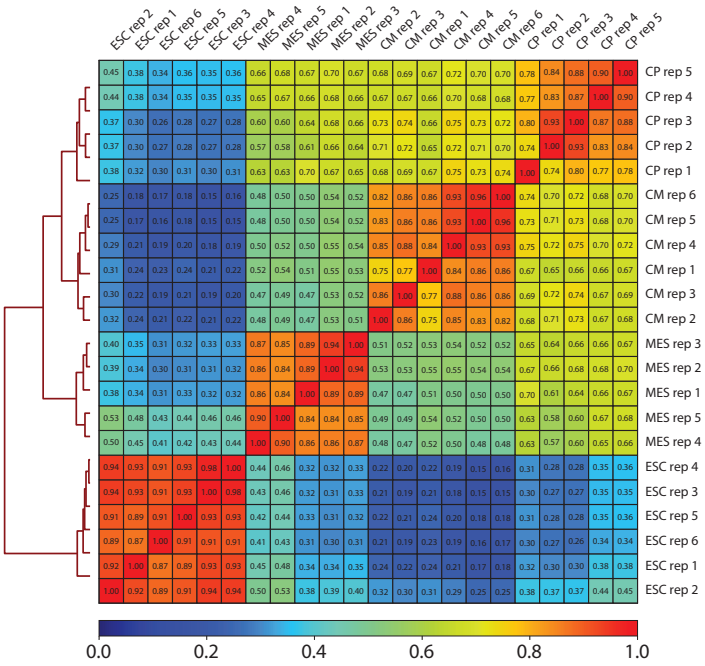


Figure 2: Correlation Between ATAC-seq Replicates

Heatmap and clustergram of correlations of ATAC-seq replicates across all open chromatin regions. Correlations are expressed in terms of Spearman's rho using the deepTools package (Ramírez et al., 2016).

Replicates showed high concordance across open regions at each time point (**Figure 2**). We found the highest number of overall ATAC-seq peaks (97,244) as well as the highest number of cell-type-specific peaks (49,065) in ESCs, consistent with observations that pluripotent cells contain a more permissive chromatin landscape compared to differentiated cells (Efroni et al., 2008; Grigoryev et al., 2006; Meshorer et al., 2006; Stergachis et al., 2013).

Because changes in chromatin accessibility during CM differentiation are expected to reveal binding of developmentally linked TFs in an unbiased manner, we further analyzed open chromatin regions. Using our previous chromatin data sets in CM differentiation (Wamstad et al., 2012), we classified chromatin states across the genome using ChromHMM (Ernst et al.,

2011). This algorithm uses multiple inputs that include histone modifications and RNA polymerase II (RNAPII) occupancy allowing for refined predictions about chromatin states across the genome (**Figure 3A**). Using the genomic annotations from ChromHMM, we observed that the majority of open chromatin peaks across CM differentiation overlapped with either active promoters or enhancers (poised or active) (**Figure 3B**). Further analysis revealed that stage-specific ATAC-seq peaks are substantially enriched for active or poised enhancers, whereas static peaks (found in all 4 stages) are largely enriched at active promoters. We also observed a dramatic increase in stage-specific active enhancer peaks compared to earlier stages, consistent with the committed state of these cells. These data support the idea that most promoters retain a nucleosome-depleted region throughout development (static open chromatin), whereas tissue-specific enhancers show dramatic differences in chromatin accessibility concomitant with TF occupancy (Thurman et al., 2012). Taken together, these analyses indicate the high quality of our data and suggest that using chromatin dynamics can identify unique sets of TFs controlling developmental gene expression programs.

Our ChromHMM annotations distinguish between two types of "enhancer" states – the "active/strong enhancer" state which displays a higher enrichment of H3K27ac, H3K4me1 as well as Serine 5-phosphorylated RNA polymerase II (RNAPSer5) compared to "poised/weak enhancers" enriched mainly for H3K4me1 alone. RNAPSer5 is a marker of transcriptional initiation and active transcription has been associated with higher enhancer activity (Zhu et al., 2013). Importantly, active enhancers classified by ChromHMM have been empirically shown using reporter assays to have cell-type-specific activity as compared to active enhancers in other cell types or poised enhancers in the same cell type (Ernst et al., 2011). As such, we identified 190,955 putative active enhancers across all four stages of CM differentiation (**Figure 4**). As expected, the genes in proximity to cell-type-specific enhancers are related to their cell identity. Thus, these putative regulatory elements likely represent regions bound by lineage-specific TFs to regulate cell-type-specific gene expression.

A

H3K27me3	H3K4me1	H3K27ac	H3K4me3	RNAPSer5	
7.6	0.1	0.6	0.7	2.0	low signal
50.9	1.7	1.5	0.2	5.7	Polycomb repressed
87.1	92.5	22.5	44.2	14.6	bivalent promoter
9.4	59.9	13.5	0.8	5.0	weak/poised enhancer
6.3	99.7	76.9	23.6	15.0	strong/active enhancer
4.3	85.2	38.0	4.2	71.7	transcriptional termination
4.6	1.2	4.7	0.3	42.0	active transcription
6.1	67.1	87.0	97.9	70.2	active promoter

B

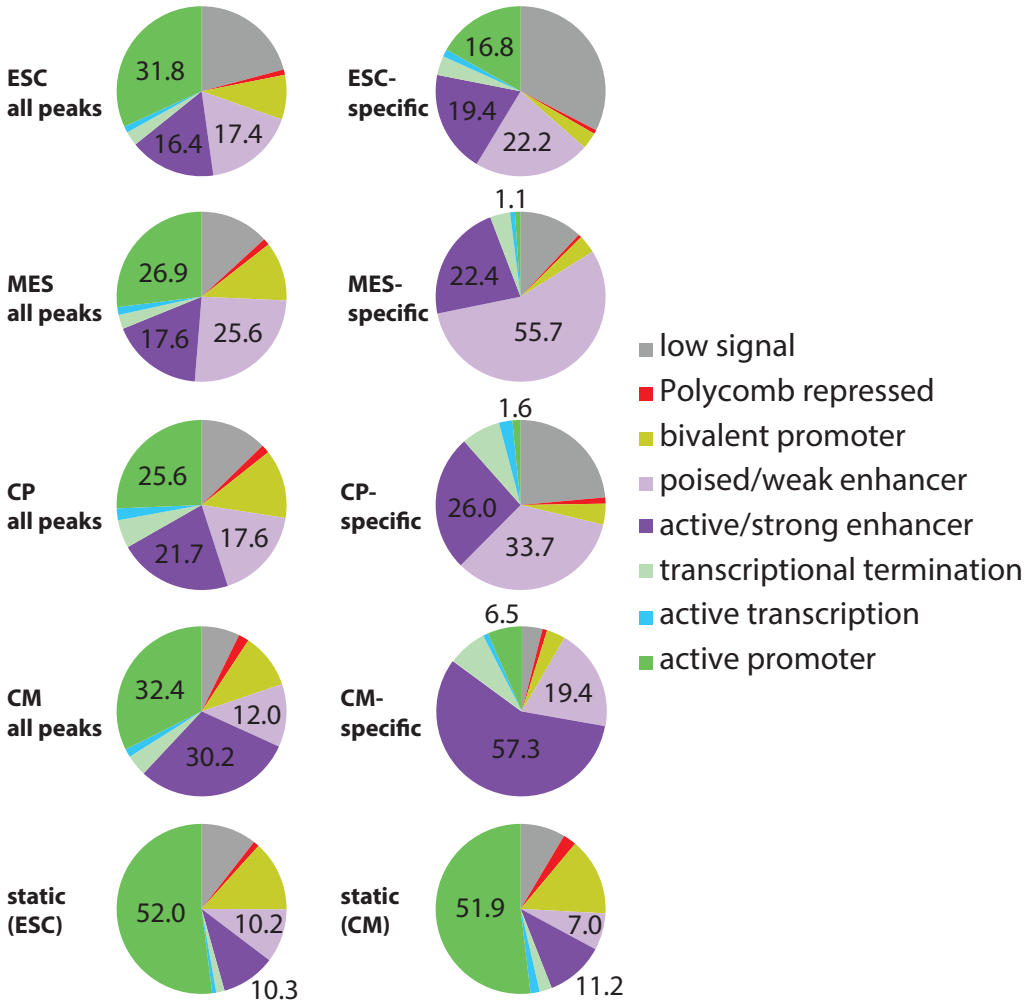


Figure 3: Identification of Chromatin States in CM Differentiation

A) Chromatin state classification using ChromHMM (Ernst et al., 2011). ChIP-seq data for histone modifications (H3K27me1, H3K4me1, H3K27ac, H3K4me3, and H3K36me3) and transcribing RNA polymerase (RNAPSer5) taken from Wamstad et al. (2012). Numbers correspond to the frequency of observation of the modification by percentage. Classifications on the right chosen based on previous associations of chromatin states with histone modifications.

B) Pie charts showing distribution of chromatin states in all ATAC-seq peaks by stage and selected subsets. Percentages for poised/weak enhancers, active/strong enhancers, and active promoters are shown.

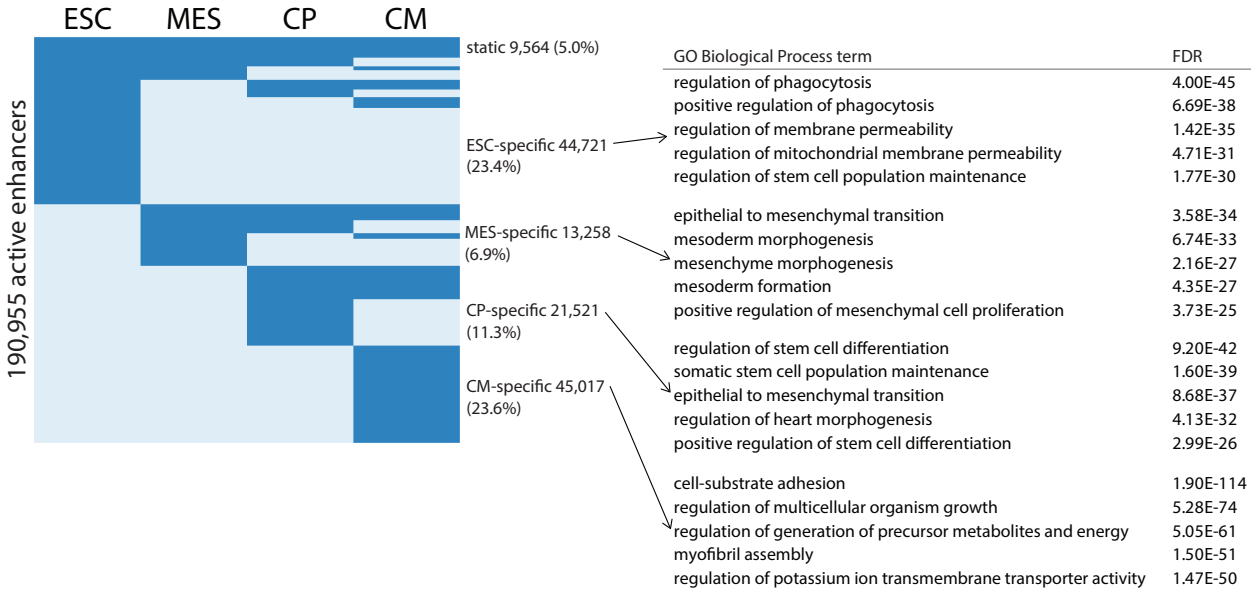


Figure 4: Active Enhancers Mark Sites of Cell-Type-Specific Gene Regulation
Heatmap of active enhancers in CM differentiation annotated by ChromHMM. Gene Ontology enrichment for biological process categories of genes in proximity to stage-specific enhancers is shown using GREAT (McLean et al., 2010). Top five terms for each stage are reported with binomial FDR.

TEAD family motifs are enriched at dynamic open chromatin regions within active enhancers

Studies suggest that clusters of cardiac TF binding sites in fetal and adult hearts can identify *bona fide* enhancers in the absence of active enhancer histone marks; however, the presence of both TF binding sites and active enhancer marks is a more reliable prognostic of enhancer activity *in vivo* (Akerberg et al., 2019). Given that lineage TFs bind densely to tissue-specific enhancers, we next determined TF motif enrichment within stage-specific open chromatin in active enhancers using the motif-enrichment algorithm in the HOMER suite (Heinz et al., 2010) (**Figure 5A**). Motifs associated with Oct, Klf, Sox, Nanog, and Esrrb TFs were highly enriched in ESCs, consistent with their roles as master regulators of pluripotency and somatic cell reprogramming (Adachi et al., 2018; Boyer et al., 2005; Festuccia et al., 2017; Takahashi and Yamanaka, 2006). In MES, our analysis identified a motif matching the T-box family of factors, which includes EOMES, a factor crucial for mesendoderm specification (Costello et al., 2011). We also identified motifs associated with Gata, Lhx, and Zic TFs, factors with known roles in mesendoderm-specific gene regulation (Costello et al., 2015; Peterkin, 2003; Winata et al.,

2013). An unspecified basic helix-loop-helix (bHLH) motif was found to be enriched in the MES-specific enhancer open chromatin. Although we could not determine which factor matched this motif, bHLH TFs are broadly implicated in regulating signal-dependent gene expression (Gyoja, 2017). As expected, a Gata motif was also highly enriched in both CP and CM-specific open chromatin, consistent with roles for this TF family in heart development (Afouda et al., 2008; Luna-Zurita et al., 2016; Wamstad et al., 2012). The enrichment of the Fox (Forkhead) motif in CP-specific open chromatin agrees with the known role of Forkhead factors in CP specification (Ahmad et al., 2015; Kang et al., 2009). In addition, the motif for the Mef2 TF family, master regulators of skeletal and cardiac muscle gene expression (Chambers et al., 1994; Edmondson et al., 1994; Potthoff and Olson, 2007), is most enriched in CM-specific open chromatin. We also identified a match for a known motif for the transcription factor MEIS1 which also closely matches a motif that has been shown to be bound by the cardiac TF TBX5 in CMs (Luna-Zurita et al., 2016).

In addition to these well-studied cardiac TFs, we identified the Tead and AP-1 motifs enriched in CP- and CM-specific open chromatin. The fact that these motifs for these TF families along with the Gata motif, are enriched in CP- and CM-specific enhancer open chromatin suggests that these factors may be driving distinct cardiac gene expression programs. Linkages between TEAD factors and the AP-1 family have been demonstrated in other studies (Vierbuchen et al., 2017; Zanconato et al., 2015), although neither of these TFs nor their potential partnership has been well-characterized in cardiac development. Thus, our data demonstrate the ability of this approach to uncover key regulatory TFs and TF linkages without *a priori* knowledge of their functions.

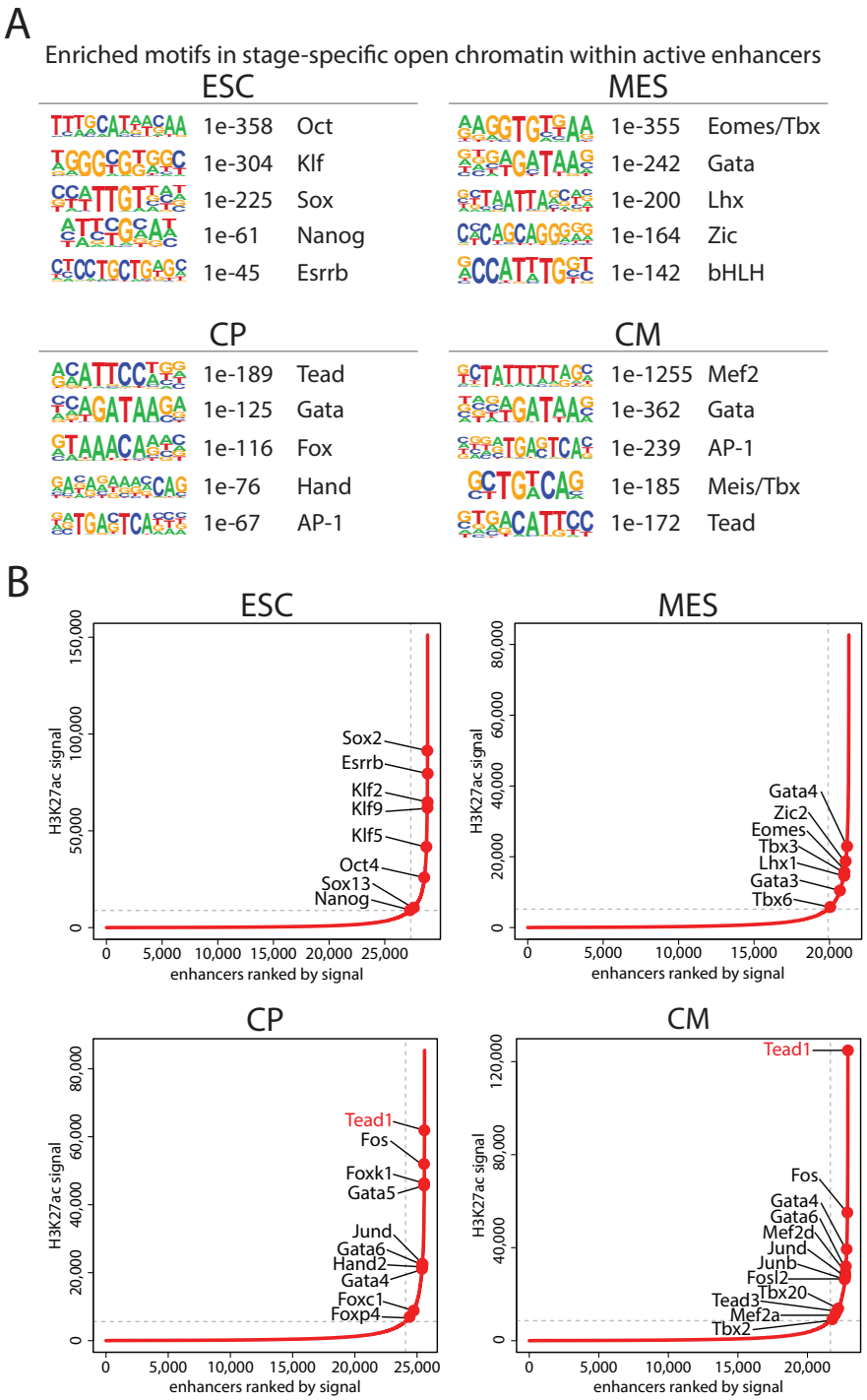


Figure 5: Identification of Candidate Master Transcription Factors

A) Enriched motifs in stage-specific open chromatin within active enhancers, identified using HOMER (Heinz et al., 2010). Motif, closest matching TF or TF family, and p-value are listed at each stage with top 5 significant motifs. **B)** Plots showing enhancers ranked in ascending order by H3K27ac signal generated using the ROSE algorithm (Lovén et al., 2013; Whyte et al., 2013). Super-enhancers are to the right of the dotted line. Displayed are super-enhancers overlapping or proximal to a TSS corresponding to a stage-specific TF in each stage matching the motif predictions identified in **(A)** as indicated by red dots.

We next sought to prioritize potential regulators of cardiac lineage commitment. Super-enhancers are *cis*-regulatory elements that consist of large enhancer clusters that often overlap or are in close proximity to genes encoding "master regulator" TFs central to cell identity (Hnisz et al., 2013; Whyte et al., 2013). To this end, we mapped super-enhancers in all 4 stages by using the ROSE algorithm (Lovén et al., 2013; Whyte et al., 2013) and determined the closest transcription start site (TSS) (**Figure 5B**). Consistent with the idea that master TFs auto-regulate their own expression, we found that a prominent super-enhancer overlaps the *Tead1* TSS in both CP and CM, suggesting that TEAD1 has crucial roles in regulating cardiac cell identity (**Figure 5B, lower panels**). Prior studies reported that specific Tead factors can regulate CM-specific genes *in vivo* (Azakie et al., 2005; Chen et al., 2004b, 2004a; Maeda et al., 2002), although *Tead1* is the only family member for which knockouts show a cardiac phenotype (Chen et al., 1994; Kaneko et al., 2007; Nishioka et al., 2008; Sawada et al., 2008). *Tead1*-knockout mice are embryonic lethal between E11 and E12 and show cardiovascular defects (Chen et al., 1994). Moreover, TEAD1 co-occupies putative enhancers in fetal and adult mouse hearts and appears to bind in combination with six other cardiac TFs (GATA4, MEF2A, MEF2C, NKX2-5, SRF, and TBX5), although its role in regulating cardiac lineage commitment was not examined (Akerberg et al., 2019). Consistent with a potential role in cardiac lineage commitment, we found that *Tead1* expression is highest in CP and CM stages at both the transcript and the protein level compared to other family members (**Figure 6A-B**). Together, these data suggest that *Tead1* occupies an important position in the TF hierarchy that regulates heart development.

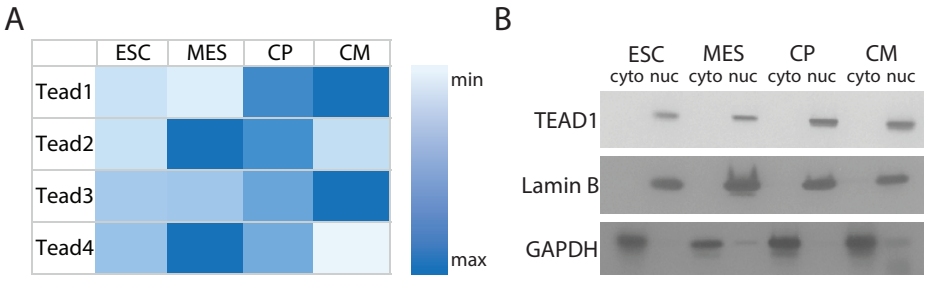


Figure 6: TEAD1 is Expressed Specifically in Cardiac-Specific Cell Types
A) Heatmap showing TEAD TF RNA-seq expression data across CM differentiation (Wamstad et al., 2012). Expression is row-normalized, meaning that each gene's expression is scaled by the maximum value across CM differentiation.
B) Immunoblot of TEAD1 across CM differentiation. Lamin B (nuclear) and GAPDH (cytoplasmic) are included as loading controls.

TEAD1 is necessary to maintain normal beating patterns in CMs

The role of TEAD1 during heart development is poorly understood relative to other cardiac TFs. Thus, we generated a homozygous ESC line that harbors a modified FKBP12 domain ($FKBP12^{F36V}$) and a double hemagglutinin tag (2xHA) at the 3' end of the endogenous Tead1 gene to study its role in our system (**Figure 7A, B**). Importantly, fusion with the $FKBP12^{F36V}$ domain allows for rapid (within 4 hours; data not shown) and near complete TEAD1 depletion upon the addition of the small molecule dTAG-13 (dTAG) (Nabet et al., 2018; Weintraub et al., 2017) (**Figure 7C**).

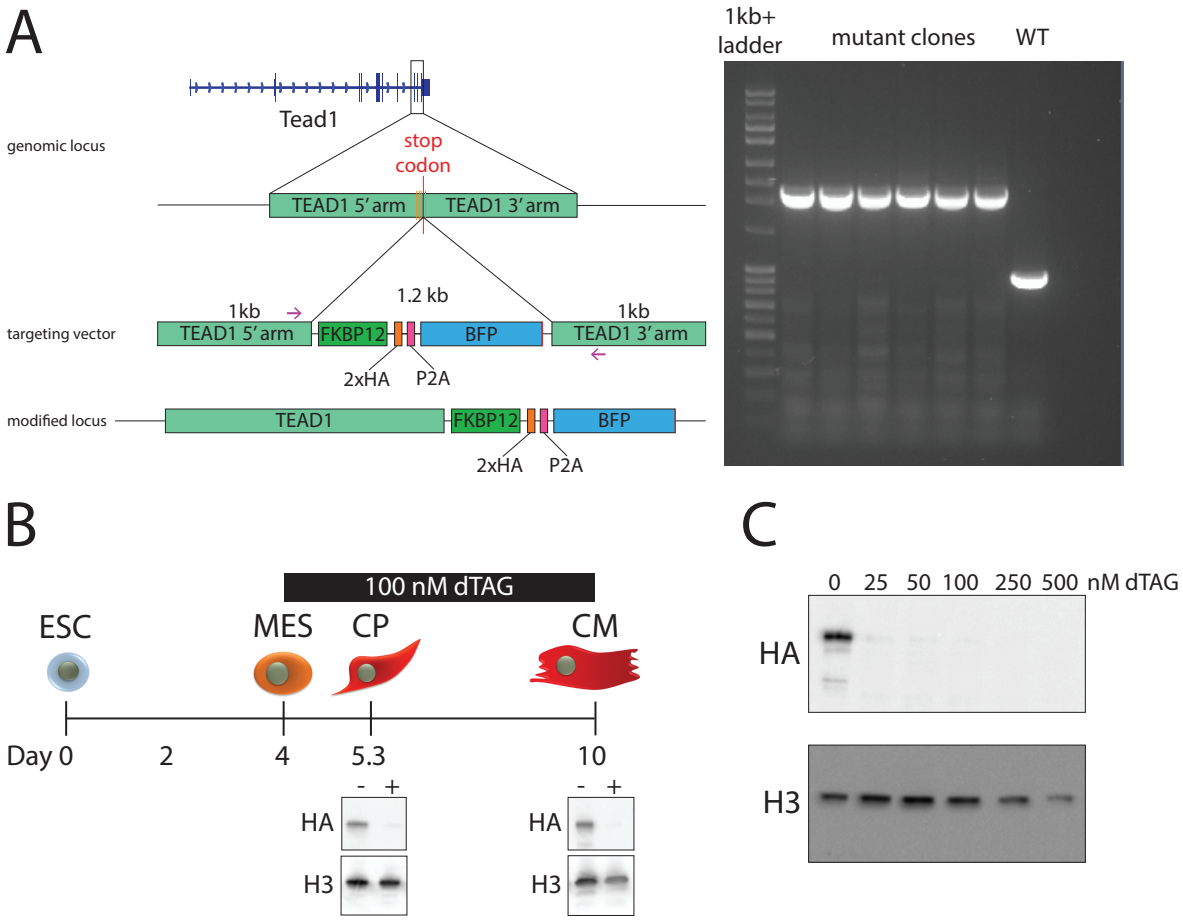


Figure 7: dTAG System Allows Rapid Depletion of TEAD1

A) Schematic showing CRISPR-mediated insertion of $FKBP12^{F36V}$ domain, and primers used for PCR genotyping (purple arrows). Guide RNA sites are indicated with orange bars. At right, gel electrophoresis of a genotyping PCR is shown.

B) CM differentiation schematic showing addition of dTAG on top and immunoblot of TEAD1 using an anti-HA antibody with 100nM dTAG treatment at CP and CM timepoints compared to controls. Histone H3 is included as a loading control.

C) Immunoblot showing TEAD1-HA degradation after 6 hours in mESCs using anti-HA antibody after addition of varying concentrations of dTAG. Histone H3 included as a loading control.

To test the function of TEAD1 during CM differentiation, we added 100 nM dTAG to MES stage cultures (day 4), a time point prior to its increased expression based on our RNA-seq and immunoblot data (**Figure 6**). We observed substantial depletion of Tead1 in CPs and complete loss in CMs by HA immunoblot compared to untreated controls (**Figure 7B**). Surprisingly, we did not observe significant changes in the proportion of Nkx2.5- or cTnT-expressing cells by FACS in dTAG-treated cells, general markers of CP and CM states, relative to untreated controls (**Figure 8A**). Upon closer examination, however, TEAD1-depleted CMs consistently demonstrated an irregular beating pattern (**Figure 8B** and see linked videos). This observation suggested that TEAD1 loss affects CM function rather than lineage commitment *per se*, a surprising result given the specific enrichment of Tead motifs in CM-specific open chromatin in enhancers.

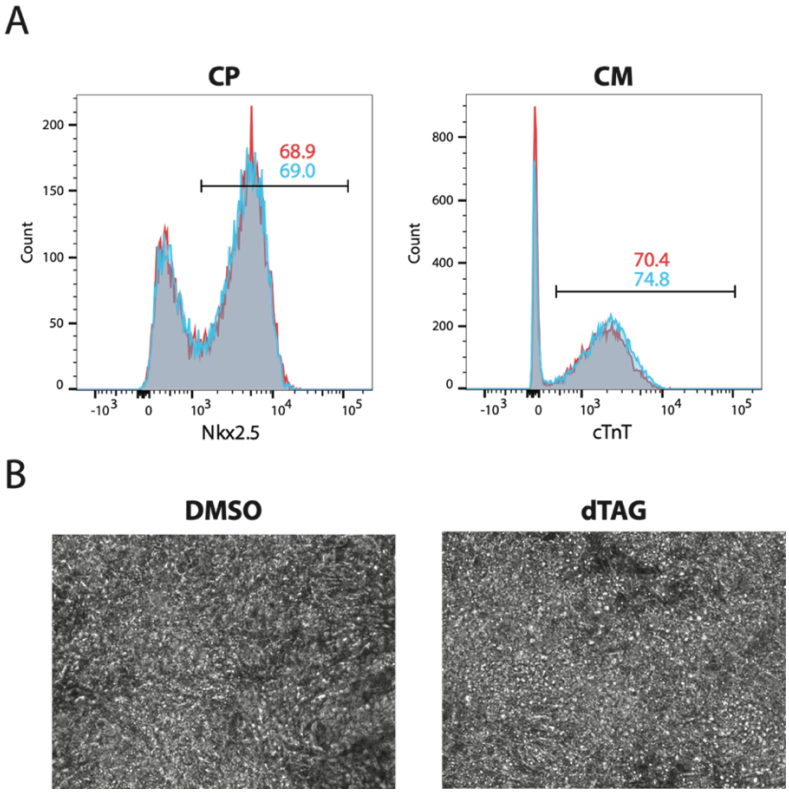


Figure 8: TEAD1 Depletion Leads to an Abnormal Beating Phenotype
A) Plots showing the distribution of Nkx2-5+ (CP) and cTnT+ (CM) cells in untreated (DMSO) controls compared to dTAG-treated cells. DMSO-treated cells are in red and dTAG-treated in blue.
B) Representative images of videos of untreated (DMSO) and dTAG-treated CMs taken on day 10 of the CM differentiation timecourse. Videos from three biological replicates can be found at https://www.dropbox.com/sh/fvgtwrpnrf323om/AADUmPvFXIKGH_colkBd7C6Ha?dl=0: 1-4 from replicate 1, 5-6 from replicate 2, and 7 from replicate 3.

TEAD1 binds to active cardiomyocyte enhancers

We next asked if TEAD1 directly contributes to enhancer activation during CM differentiation. To this end, we analyzed available ChIP-seq data in fetal and adult mouse heart samples from mice expressing biotinylated TFs including GATA4, NKX2-5, MEF2A, MEF2C, SRF, TBX5, and TEAD1 (Akerberg et al., 2019). We first compared ATAC-seq data from fetal (E12.5) and adult heart (P56) samples (Quaife-Ryan et al., 2017) to our *in vitro* differentiated CM data. We found high concordance between the accessible chromatin landscape in fetal heart and *in vitro* differentiated CMs (ρ = 0.8) relative to adult heart (ρ = 0.69) (**Figure 9A-C**). Notably, CMs differentiated from pluripotent stem cells are highly similar in gene expression to CMs in early to mid-embryonic hearts (Uosaki et al., 2015). Moreover, comparisons of H3K27ac signal also showed high concordance between *in vitro*-differentiated CMs and fetal hearts (ρ = 0.7) as compared to adult hearts (ρ = 0.58) (Nord et al., 2013) (**Figure 9D-F**). These data suggest that the binding of cell-type-specific TFs in fetal heart is likely representative of that in our *in vitro*-differentiated CMs and consistent with the notion that *in vitro*-derived cell types are less mature.

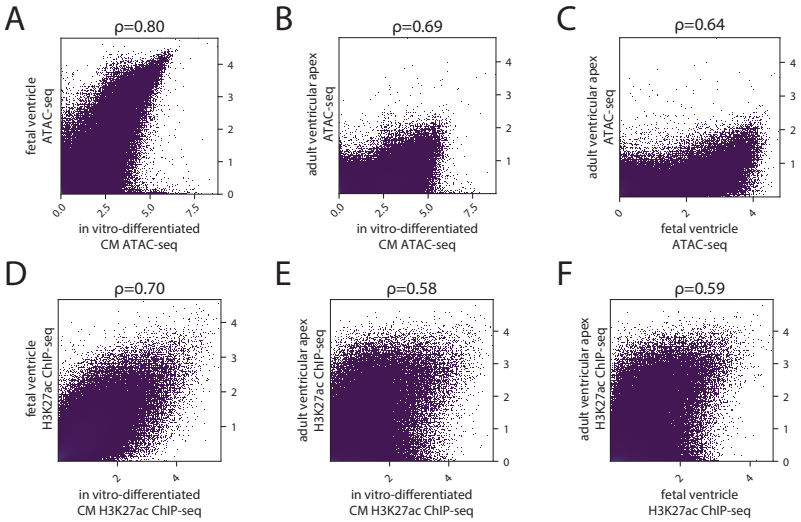
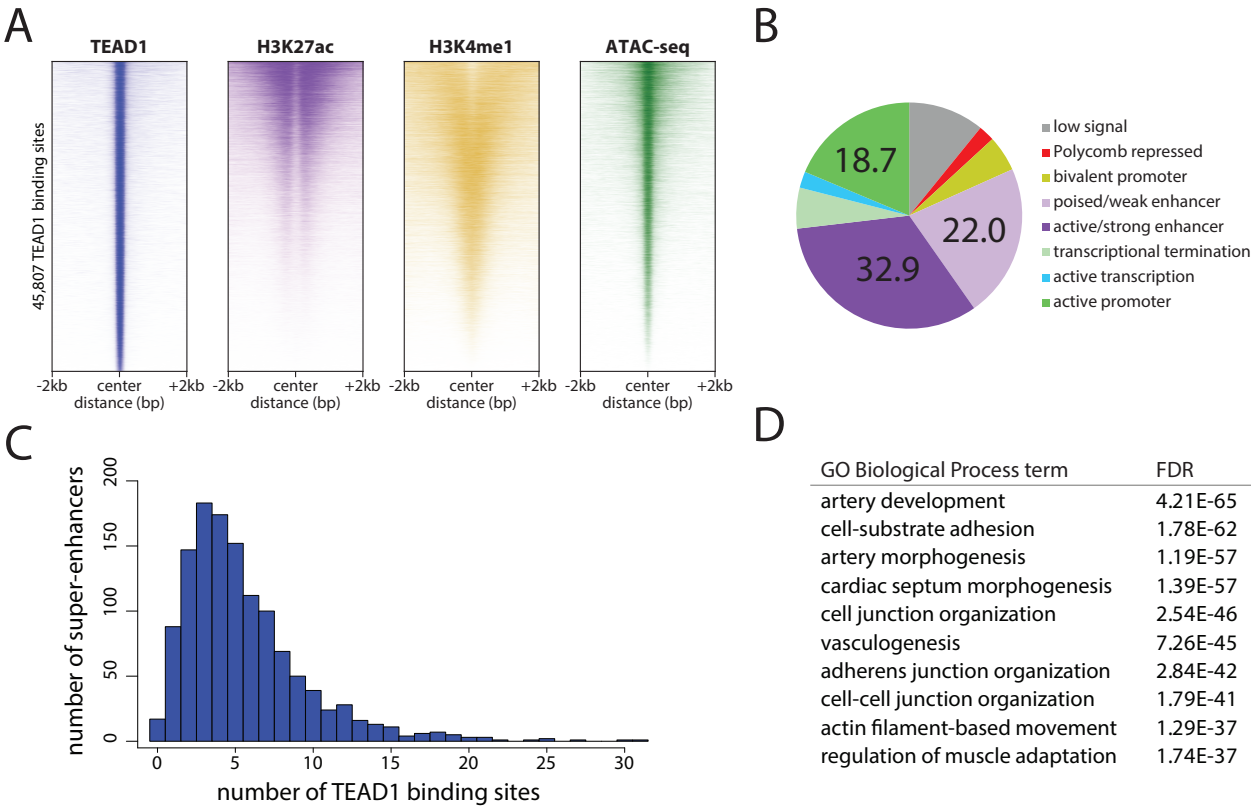


Figure 9: Accessible Chromatin and Active Enhancer Landscapes in Fetal Heart Correlate with *In Vitro*-Differentiated CMs

Scatterplots comparing ATAC-seq experiments from fetal ventricle (E12.5) **(A)** and adult heart (P56) **(B)** with *in vitro*-differentiated CMs, as well as to each other **(C)**, across the union of ATAC-seq peaks found in all 4 stages of *in vitro* CM differentiation. The same analysis for H3K27ac signal, across the union of active enhancers, is shown in **(D)**, **(E)**, **(F)**. The Spearman correlation coefficient (ρ) is reported at the top of each plot.

We next analyzed the 45,807 TEAD1 binding sites identified in the ventricles of fetal hearts (Akerberg et al., 2019), a region known to mostly consist of CMs (**Figure 10A**). The binding density at these sites correlated strongly with histone modifications associated with enhancer activity such as H3K27ac and H3K4me1 as well as chromatin accessibility as measured by ATAC-seq in *in vitro* CMs; these correlations are also found with fetal ventricular H3K27ac and ATAC-seq data (data not shown). In fact, the majority of TEAD1 binding sites from fetal heart (54.9%) are found within putative CM enhancers (**Figure 10B**). Moreover, nearly all of the 1,258 super-enhancers identified by H3K27ac signal in the *in vitro*-derived CMs had at least one TEAD1 binding site with a substantial number containing multiple binding sites (**Figure 10C**). These data are consistent with our analysis showing that Tead motifs are highly enriched in CM-enhancer open chromatin (**Figure 5A**) and strongly suggest that TEAD1 has critical functions in cardiac lineage commitment. Thus, we next analyzed the putative target genes nearest to TEAD1-bound active enhancers, and found enrichment for Gene Ontology terms related to vascular development, cell-cell junctions, and actin cytoskeletal structure using GREAT (McLean et al., 2010) (**Figure 10D**). These categories are consistent with the abnormal beating phenotype observed upon dTAG treatment, suggesting that TEAD1 binds to CM enhancers and regulates genes with roles in cell-cell contacts and cardiac contraction.



GO Biological Process term	FDR
artery development	4.21E-65
cell-substrate adhesion	1.78E-62
artery morphogenesis	1.19E-57
cardiac septum morphogenesis	1.39E-57
cell junction organization	2.54E-46
vasculogenesis	7.26E-45
adherens junction organization	2.84E-42
cell-cell junction organization	1.79E-41
actin filament-based movement	1.29E-37
regulation of muscle adaptation	1.74E-37

Figure 10: TEAD1 Binds a Subset of Active Enhancers in CMs

A) Heatmap of TEAD1 peaks in fetal ventricles compared to CM H3K27ac and H3K4me1 from our earlier study (Wamstad et al., 2012), and ATAC-seq from this study. Heatmap is sorted by TEAD1 binding intensity.
B) Pie chart showing chromatin states of TEAD1 binding sites in CM.
C) Histogram showing frequency of TEAD1 binding sites within super-enhancers.
D) Gene Ontology enrichment for biological process and cellular component categories of genes in proximity to TEAD1-bound enhancers is shown using GREAT (McLean et al., 2010). Top 10 terms are reported with binomial FDR.

TEAD1 regulates enhancer activity during cardiac lineage commitment

If TEAD1 acts as a key regulator, we expected that its depletion during CM differentiation also lead to a decrease in chromatin accessibility at TEAD1 enhancer binding sites. To test this idea, we performed ATAC-seq in dTAG-treated and untreated CMs. ChIP-seq experiments using an anti-HA antibody showed a dramatic loss of TEAD1 at binding sites upon dTAG treatment in the TEAD1-FKBP line (**Figure 11**).

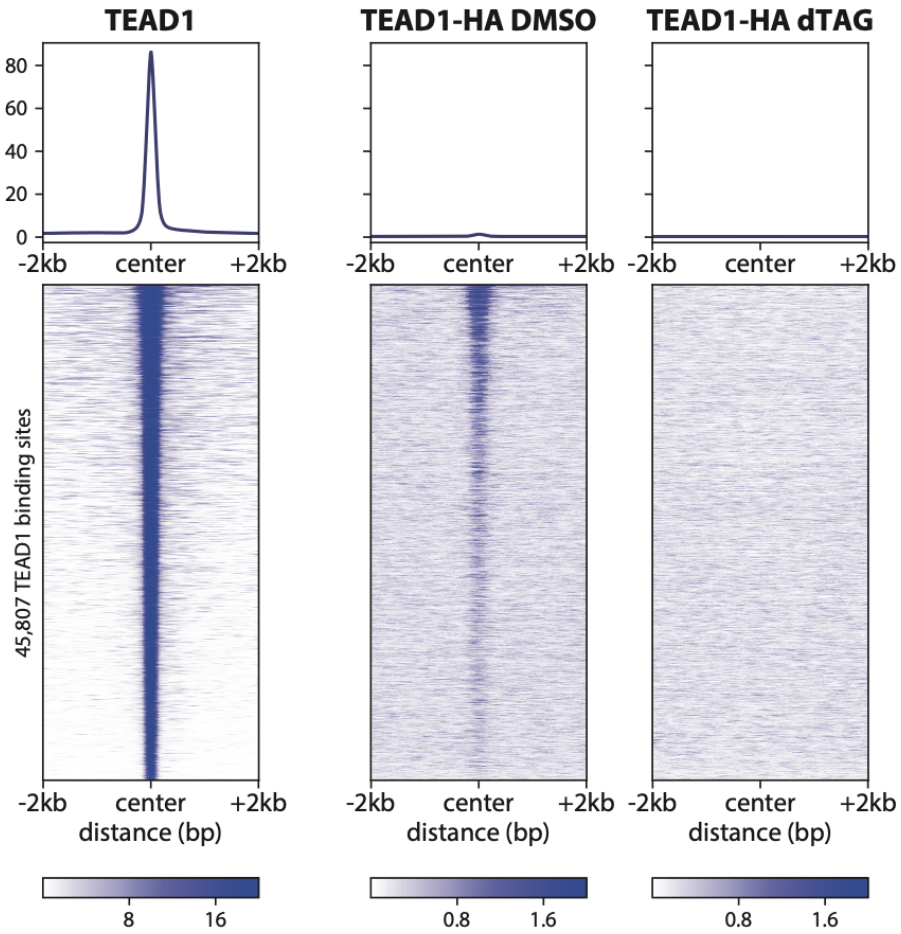
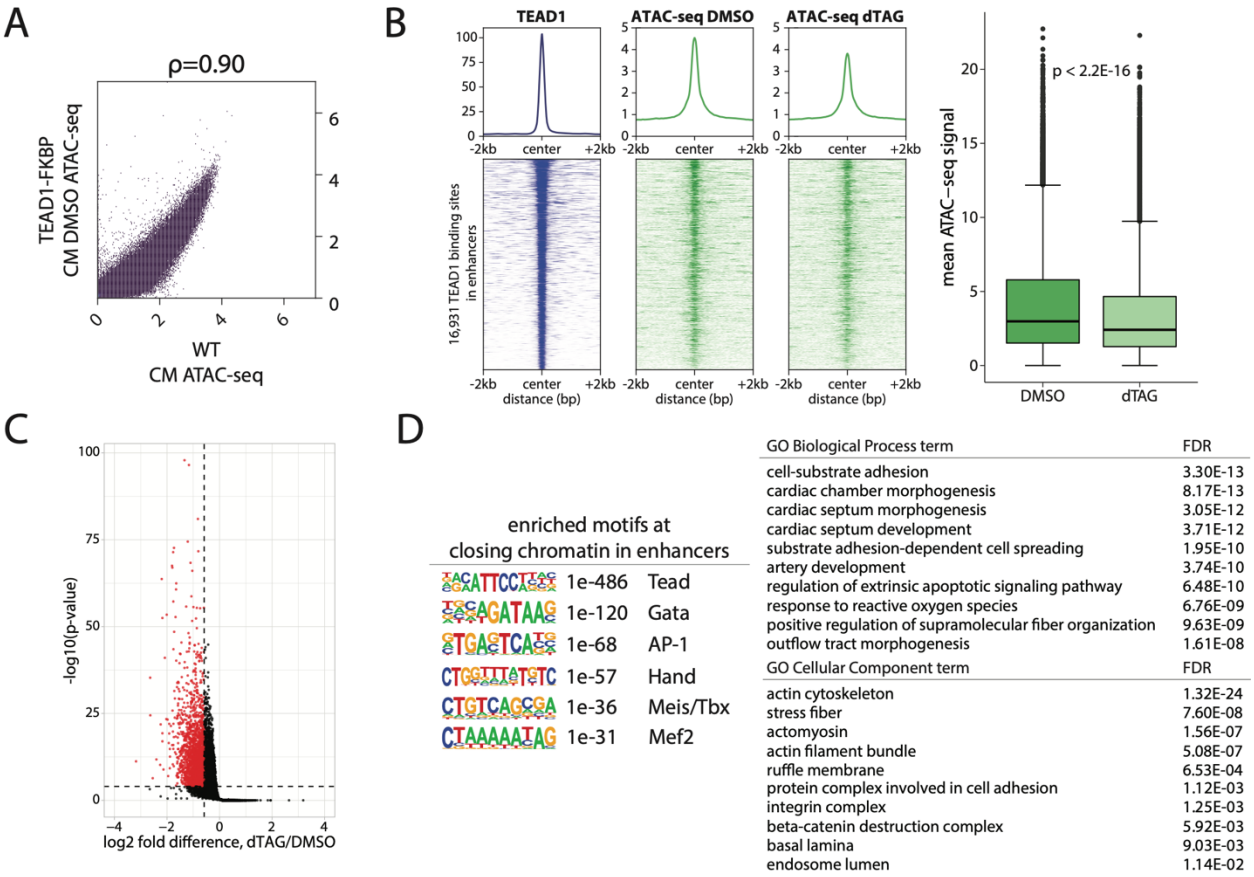


Figure 11: TEAD1 is Depleted Genome-Wide by dTAG Treatment

Heatmap showing TEAD1 depletion at binding sites by treatment with 100 nM dTAG as depicted in Figure 7B. Left to right, TEAD1 ChIP-seq data from fetal ventricles (Akerberg et al., 2019), TEAD1-HA ChIP-seq data from *in vitro*-differentiated TEAD1-FKBP CMs treated with DMSO (vehicle), and TEAD1-HA ChIP-seq from *in vitro*-differentiated TEAD1-FKBP CMs treated with dTAG. Heatmap intensity scale is shown below each panel.

We found a substantial and significant decrease in open chromatin at TEAD1 binding sites within active enhancers in dTAG-treated cells compared to untreated cells by ATAC-seq (**Figure 12B**). Analysis of DMSO-treated controls showed high concordance with our initial ATAC-seq data obtained from parental ESC lines, indicating that DMSO treatment or targeting of these cells does not affect our results (**Figure 12A**). We next identified ATAC-seq peaks that showed the most significant decrease in chromatin accessibility upon dTAG treatment. Using a Poisson distribution-based approach (Heinz et al., 2010), we identified 2,902 open chromatin peaks ins active enhancers that showed at least a one-third decrease in accessibility and were statistically significant (p-value < 1E-4) (**Figure 12C**). Over two-thirds of these regions (1,973; 68.0%) were directly bound by TEAD1 based on fetal ventricle ChIP-seq, suggesting that a substantial

proportion are directly affected by loss of TEAD1 binding. Motif enrichment analysis showed that motifs corresponding to other cardiac TFs were most enriched in these regions (**Figure 12D**), suggesting that TEAD1 binding is necessary for the cooperative binding of other factors.



GO Biological Process term	FDR
cell-substrate adhesion	3.30E-13
cardiac chamber morphogenesis	8.17E-13
cardiac septum morphogenesis	3.05E-12
cardiac septum development	3.71E-12
substrate adhesion-dependent cell spreading	1.95E-10
artery development	3.74E-10
regulation of extrinsic apoptotic signaling pathway	6.48E-10
response to reactive oxygen species	6.76E-09
positive regulation of supramolecular fiber organization	9.63E-09
outflow tract morphogenesis	1.61E-08
GO Cellular Component term	**FDR**
actin cytoskeleton	1.32E-24
stress fiber	7.60E-08
actomyosin	1.56E-07
actin filament bundle	5.08E-07
ruffle membrane	6.53E-04
protein complex involved in cell adhesion	1.12E-03
integrin complex	1.25E-03
beta-catenin destruction complex	5.92E-03
basal lamina	9.03E-03
endosome lumen	1.14E-02

Figure 12: Chromatin Accessibility Decreases at TEAD1-Bound Enhancers

A) Correlation scatterplot of ATAC-seq data between vehicle-treated (DMSO) CMs and untreated CMs, across all ATAC-seq peaks as in Figure 9A-C. The Spearman correlation coefficient (ρ) is reported at the top of the plot.

B) Left, heatmap and profile showing decrease in ATAC-seq signal at TEAD1 binding sites in active enhancers. Fetal ventricle TEAD1 binding data is shown on left with ATAC-seq performed on DMSO-treated (middle) and dTAG-treated (right) *in vitro*-differentiated CMs. At right, boxplot showing the change in distribution of mean ATAC-seq signal at TEAD1 binding sites in active enhancers between DMSO-treated and dTAG-treated cells. Significance was determined by Mann-Whitney *U* test.

C) Volcano plot of ATAC-seq signal changes under dTAG treatment at open chromatin peaks within enhancers. Horizontal axis is the log2 fold difference between dTAG-treated cells and DMSO-treated cells, and the vertical axis is the negative log10 p-value calculated based on a Poisson distribution (Heinz et al., 2010). Dotted lines mark the significance cutoffs – 1.5-fold difference and 1E-4 p-value, and significantly decreased peaks are marked in red.

D) Left, enriched motifs in open chromatin regions that showed reduced accessibility upon dTAG treatment identified using HOMER (Heinz et al., 2010). Motif, closest matching TF or TF family, and p-value are listed with top six significant motifs. Right, Gene Ontology enrichment for biological process and cellular component categories of genes in proximity to those regions is shown, using GREAT (McLean et al., 2010). Top five terms are reported for each along with their binomial FDR.

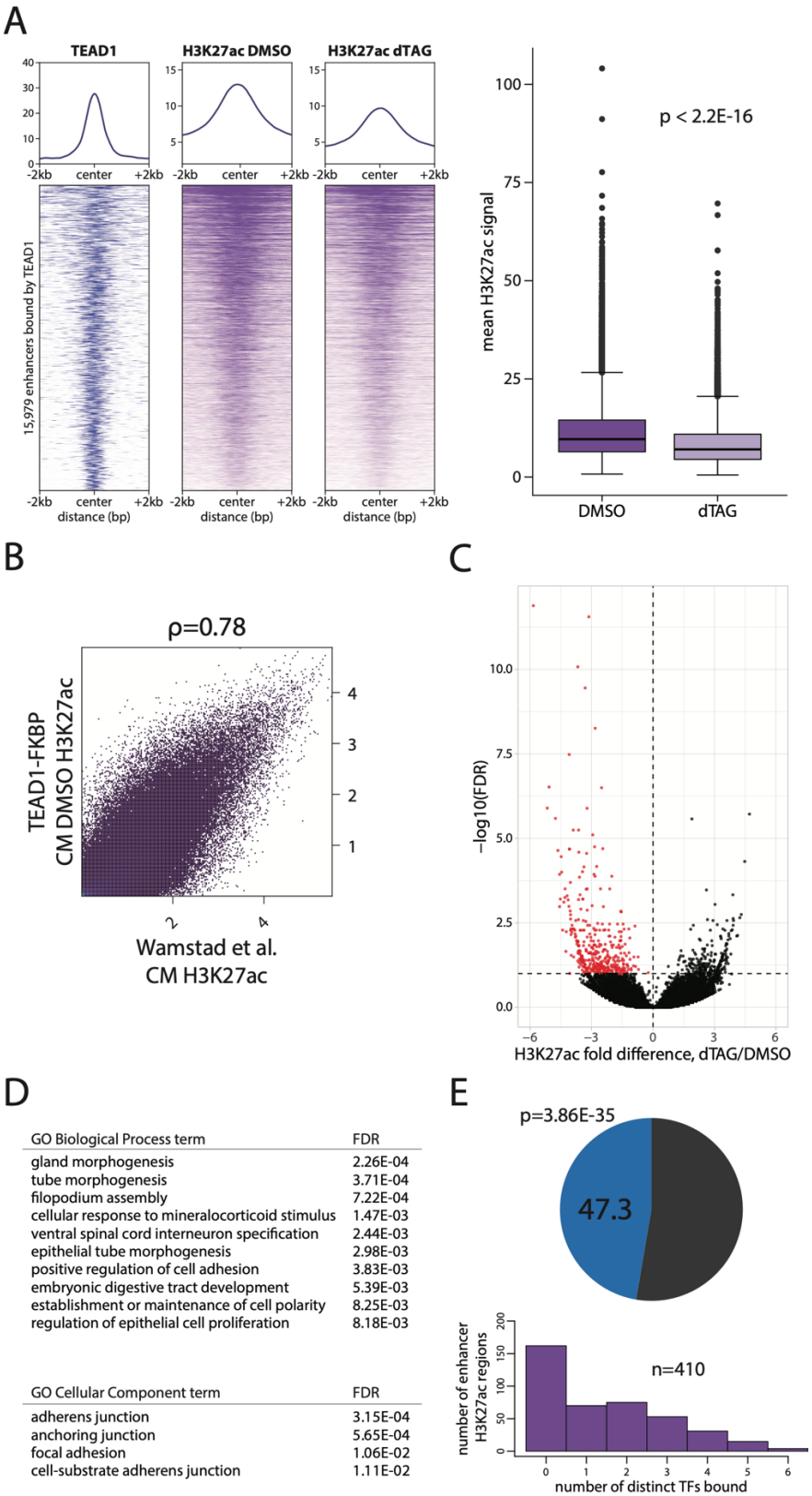


Figure 13: TEAD1 Depletion Leads to Loss of Active Enhancer Marks

A) Left, heatmap and profile showing decrease in H3K27ac at TEAD1-bound active enhancers, centered on the middle of the enhancer region annotated by ChromHMM. Fetal ventricle TEAD1 binding data is on the left with H3K27ac ChIP-seq performed on DMSO-treated and dTAG-treated *in vitro*-differentiated CMs in the middle and on the right. Right, boxplot showing the change in distribution of mean H3K27ac signal at TEAD1-bound active enhancers between DMSO-treated and dTAG-treated cells, with the results of a Mann-Whitney *U* test reported on the plot.
B) Correlation scatterplot of H3K27ac ChIP-seq data between vehicle-treated (DMSO) CMs and untreated CMs (Wamstad et al., 2012), across all active enhancers as in Figure 9D-F. The Spearman correlation coefficient (ρ) is reported at the top of the plot.
C) Volcano plot of H3K27ac signal changes under dTAG treatment at H3K27ac peaks within active enhancers, as calculated by DiffBind (Ross-Innes et al., 2012). Horizontal axis is the fold difference between dTAG-treated and DMSO-treated cells, while the vertical axis is the negative log10 FDR for each peak. Dotted lines mark significance cut-offs – any decrease between DMSO and dTAG with a 10% FDR. Significantly decreased peaks are marked in red.
D) Gene Ontology enrichment for biological process and cellular component categories of genes in proximity to significantly deactivated regions in response to TEAD1 depletion denoted in **(C)** using GREAT (McLean et al., 2010). The top 10 significant terms are reported for biological process and all are reported for cellular component along with binomial FDR.
E) Top, Pie-chart indicating what percentage of significantly decreased H3K27ac enhancer regions are bound by TEAD1. P-value for a hypergeometric test comparing this overlap to all CM H3K27ac enhancers is reported. Histogram of distinct cardiac TFs binding significantly deactivated regions in response to TEAD1 depletion. TFs surveyed were TEAD1, GATA4, HAND2, TBX5, MEF2A, and MEF2C, using fetal ventricle and fetal heart data (Akerberg et al., 2019; Laurent et al., 2017).

Our data suggest that TEAD1 binding is necessary for the activation of a subset of enhancers during cardiac lineage commitment. To further test this idea, we performed ChIP-seq for H3K27ac in dTAG-treated and untreated CMs. As expected, we observed a global decrease in H3K27ac signal at TEAD1-bound active enhancers (**Figure 13A**). DMSO-treated controls in our transgenic Tead1-FKBP line were highly concordant with our previous H3K27ac ChIP-seq experiments in CMs derived from the parental cell line (**Figure 13B**) (Wamstad et al., 2012). By analyzing H3K27ac signal changes between dTAG- and DMSO-treated CMs using DiffBind (Ross-Innes et al., 2012), we found 410 H3K27ac peaks that showed a significant decrease in H3K27ac signal in active CM enhancers (**Figure 13C**). Interestingly, we did not find a direct correlation between the magnitude of fold-change differences between DMSO-treated and dTAG-treated cells in H3K27ac signal and ATAC-seq signal, even when examining the same set of genomic regions (data not shown). This observation may be partially explained by the slightly different approaches used to calculate fold-change, but it may also be an indication that changes in chromatin accessibility may represent an immediate and local effect of loss of TF binding, whereas resetting histone modification patterns represents a late event in the process.

Gene Ontology enrichment analysis of enhancers showing significant decreases upon TEAD1 loss using GREAT showed proximity to genes linked to developmental morphogenesis and cell-cell junctions, specifically adherens junctions and anchoring junctions, which include desmosomes (**Figure 13D**). These junctions are two components of the intercalated disc structure that connects mature CMs in a syncytium to allow contraction in a concerted manner to pump blood throughout the body (Vermij et al., 2017). In immature, embryonic CMs such as those produced by *in vitro* CM differentiation, the well-organized, polarized intercalated disc structure has not yet formed, but adjacent cells are still connected by adherens junctions and desmosomes (Scuderi and Butcher, 2017). Disruption of these junctions by loss of TEAD1 could have substantial phenotypic consequences and could explain the abnormal beating phenotype observed in our analysis.

If TEAD1 binding is important for enhancer activation during cardiac lineage commitment, we expect to observe a significant association between TEAD1 binding and significantly deactivated enhancers upon TEAD1 loss. About half of the enhancer H3K27ac regions that significantly decreased in signal upon dTAG treatment were bound by TEAD1 (194/410; 47.3%) (**Figure 13E**). This overlap was highly statistically significant when compared to all H3K27ac enhancer regions (p=3.86E-35). Nonetheless, the fact that a little over half of these regions were not bound by TEAD1 suggests that loss of this factor might have more indirect effects on cardiac gene regulatory networks, perhaps by reducing the expression of other TFs. As described above, TEAD1 appears to bind to the vast majority of CM super-enhancers (**Figure 10C**), suggesting that loss of this factor might affect the regulation of other cardiac TFs activated by super-enhancers.

The fact that regions of open chromatin regions that showed less accessibility upon dTAG treatment were enriched for the motifs of other cardiac TFs suggested that TEAD1 functions with other factors to regulate tissue-specific enhancers and gene expression programs (**Figure 12D**). Leveraging available binding data of cardiac TFs matching those motifs (GATA4, HAND2, TBX5, MEF2A, and MEF2C) in fetal heart (Akerberg et al., 2019; Laurent et al., 2017), we found

that many of the differential enhancers identified in our study also show binding by multiple cardiac TFs in addition to TEAD1 (**Figure 13E**). We observed several H3K27ac peaks within enhancers bound by multiple different cardiac TFs that show substantial decreases in H3K27ac signal upon dTAG treatment, suggesting that TEAD1 binding might be critical to maintaining the overall activity of these regions. Gene expression profiling will be necessary to establish a direct link between TEAD1-mediated enhancer activity and chromatin accessibility and cardiac gene regulation. For example, a CM super-enhancer close to the *Gata4* TSS is co-occupied by TEAD1, MEF2A, MEF2C, TBX5, GATA4, and HAND2, and this region shows decreased H3K27ac and open chromatin upon dTAG treatment (**Figure 14**). The fact that TEAD1 depletion appears to have such a major effect on super-enhancer putatively linked to a key cardiac TF provides further evidence that TEAD1 is an integral part of the cardiac regulatory network. Taken together, our work provides substantial evidence that TEAD1 acts as a key cardiac TF by promoting the activation of cardiac muscle enhancers and suggests that TEAD1 controls a distinct gene expression program necessary for intercellular junctions and contraction in CMs.

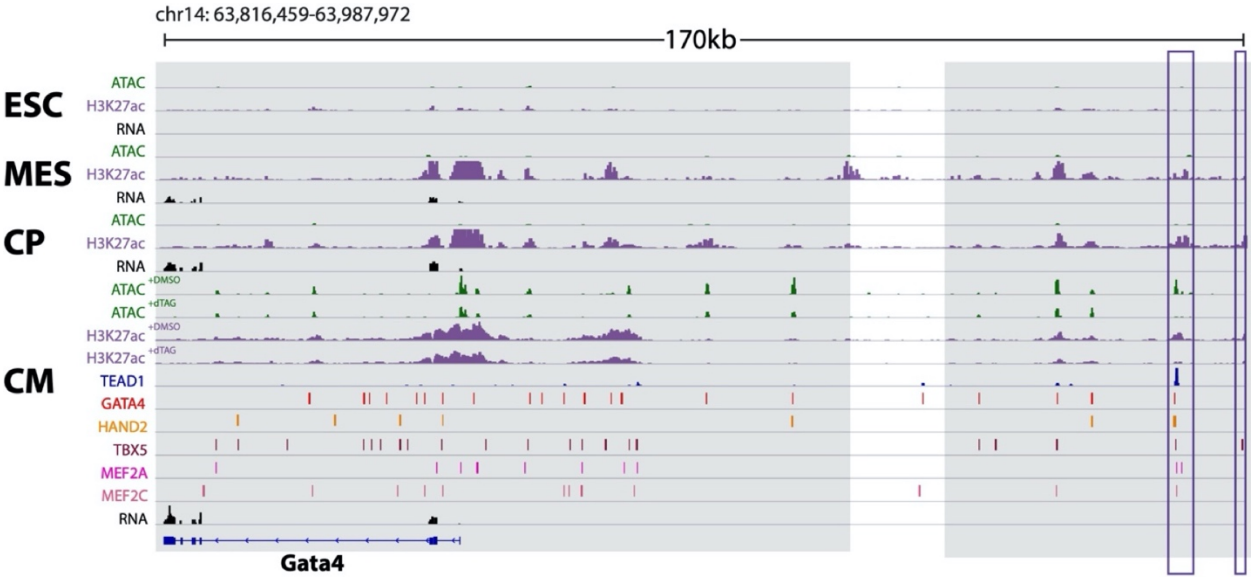


Figure 14: A *Gata4*-Linked Super Enhancer is Disrupted by TEAD1 Depletion

Tracks showing the *Gata4* locus from ESC to CM. Tracks for ATAC-seq (green) are from this study; H3K27ac (purple) from ESC to CP, and RNA-seq (black) from ESC to CM is taken from our prior CM differentiation study (Wamstad et al., 2012). For CM, ATAC-seq and H3K27ac are shown with DMSO or with dTAG treatment, as indicated. For TF binding, fetal ventricle TEAD1 (navy blue) is shown as a track, while called peaks are shown for GATA4 (red), HAND2 (orange), TBX5 (maroon), MEF2A (pink), MEF2C (light pink) (Akerberg et al., 2019; Laurent et al., 2017). Regions shaded in gray were called as super-enhancers by ROSE (Hnisz et al., 2013; Whyte et al., 2013). Regions boxed in purple were significantly reduced in H3K27ac signal as called by DiffBind (Ross-Innes et al., 2012).

DISCUSSION

We applied a systematic approach to identifying key TFs and novel linkages that regulate heart development. Our work demonstrated that the accessible chromatin landscape is highly dynamic at putative enhancers during cardiomyocyte differentiation. Through analysis of dynamic accessible chromatin in enhancers, we identified many candidate TFs that may play regulatory roles during CM differentiation. In particular, we highlight roles for the TEAD1 in controlling cardiac-specific gene expression. By combining high-throughput analysis of histone modifications and open chromatin genome-wide with protein suppression, our data suggest that TEAD1 acts to regulate tissue-specific gene expression programs distinct from other cardiac TFs. The observed phenotype of asynchronous beating suggests that TEAD1 is involved in regulating the establishment and maintenance of cell-cell connections. This idea is consistent with our finding that enhancers that showed the largest change in open chromatin and active enhancer marks upon TEAD1 depletion were in proximity to genes that code for proteins in adherens junctions and desmosomes. Further analysis on gene expression controlled by TEAD1 and the distribution of cell-cell junctions will be needed to more completely assess the role of TEAD1 in this process.

This study provides mechanistic evidence on how TFs regulate enhancer chromatin states. The fact that depletion of TEAD1 led to a significant decrease in H3K27ac at regions where many other cardiac TFs are bound suggests that at least in some cases, lineage-specific TFs are cooperative in establishing and maintaining enhancer activity. As discussed in Chapter 1, evidence suggests TFs bind at enhancers with relatively loose spatial constraints, but multiple TFs must be bound in order to control cell-type-specific activity. The data generated in this study can be leveraged to begin to address this model in context of cardiac lineage commitment. One possibility is that TEAD1 is required to establish a local chromatin environment conducive to the binding of other factors. Alternatively, TEAD1 may interact with other TFs in order to recruit them to specific loci. Evidence for the latter model would come from both a fixed binding syntax with other factors as suggested in recent studies (Akerberg et al., 2019), as well as more detailed biochemical and biophysical studies.

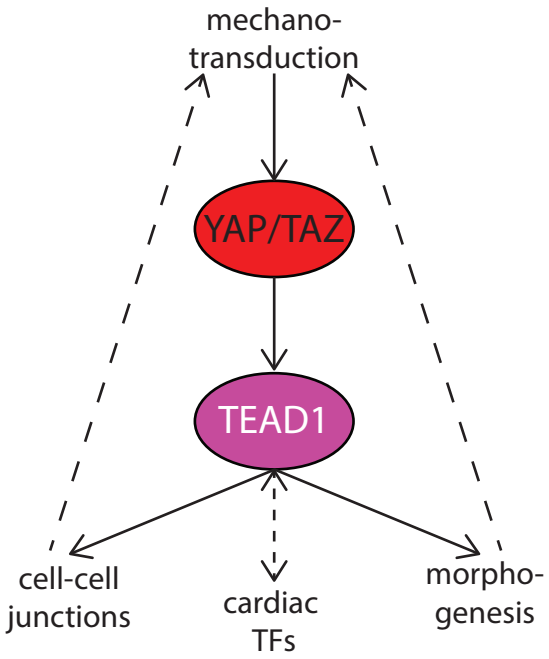


Figure 15: TEAD1 Regulates a Cardiac Gene Expression Program in Response to Stimuli
Proposed model for TEAD1-mediated gene regulation during cardiac lineage commitment. The transcriptional co-activators YAP and/or TAZ work with TEAD1 in response to mechanical cues to activate genes relating to cell-cell junctions and morphogenesis, processes that can feed back into mechanical signaling. TEAD1 may also regulate the expression of other cardiac TFs in an integrated gene network.

As for the former, it must be noted that TEAD1 is also a well-known downstream effector of extracellular signaling pathways. As discussed in Chapter 1, the Hippo kinase cascade, implicated in control of cell proliferation and organ size, regulates the nuclear entry of the co-activator YAP1 and its paralog TAZ which partner with TEAD1 to activate gene expression (Ota and Sasaki, 2008). Notably, YAP1's subcellular localization can also be regulated by mechanical cues (Chang et al., 2018; Ege et al., 2018; Elosegui-Artola et al., 2017; Mosqueira et al., 2014; Ragni et al., 2017). Studies in zebrafish provide evidence that disruption of the function of these co-activators affects morphogenesis and cell junctions, consistent with the enhancers disrupted by TEAD1 loss in our study. Expression of dominant-negative forms of Yap and Taz/Wwtr1 that couldn't interact with Tead factors blocked CPs from properly migrating to the midline, although these cells appeared to still differentiate into CMs (Miesfeld and Link, 2014). Taz/Wwtr1 knockouts in zebrafish have defects in cardiac trabeculation, a necessary process to

supply oxygen and nutrients to the developing heart, as a result of disordered cortical actin structure and irregular cell-cell junctions (Lai et al., 2018). Taken together, these data suggest a feedback mechanism in which mechanical cues related to the cellular environment promote YAP and/or TAZ nuclear entry and the activation of TEAD1-dependent gene expression programs relating to mechanical cell-cell junctions, promoting further morphogenesis in the developing heart (**Figure 15**). Empirical evidence for this mechanism could come from differentiating CMs *in vitro* on substrates of varying stiffness and observing the effect on YAP/TAZ nuclear entry and TEAD1-mediated enhancer activation, as well as blocking the LINC (Linker of Nucleoskeleton and Cytoskeleton) complex from transducing cytoskeletal signals to the nuclear envelope through dominant-negative protein expression (Elosegui-Artola et al., 2017) or disrupting the actin cytoskeleton through chemical inhibition.

Finally, understanding the interplay between extracellular signaling and chromatin states has major ramifications for understanding of how most TFs mediate gene regulation during cardiac development. As discussed above, the genetic causes of many forms of CHD remain unknown (Blue et al., 2017). Given the potential importance of TEAD1 in controlling transcriptional responses to mechanical forces as discussed above, variants linked to weakness in heart contraction or sudden cardiac death by GWAS could be intersected with TEAD1-bound enhancers, potentially providing more insight into cardiovascular disease etiology. Disruption of these enhancers, whether by loss of TEAD1 binding or through inhibiting the activity of chromatin modifiers may link enhancer-mediated gene regulation to fruitful therapeutic strategies. Overall, these insights will help us develop a deeper understanding of how lineage-specific TFs translate a variety of extracellular and intracellular stimuli into specific transcriptional outcomes during development.

MATERIALS AND METHODS

Cardiomyocyte differentiation

In vitro cardiomyocyte (CM) differentiation was performed as previously described, with some modification (Wamstad et al., 2012). Briefly, mouse embryonic stem cells (E14 Tg Nkx2-5-EmGFP) (Hsiao et al., 2008) cultured on feeders using standard technique were isolated and cultured at 100,000 cells/ml in suspension into embryoid bodies (EBs) in serum-free defined (SFD) media (3 parts Iscove's Modified Dulbecco's Medium (Millipore-Sigma #I3390), 1 part Ham's F-12 (Corning #10-080-CV), 0.05% BSA (Sigma-Aldrich #A1595), 1.5 mM L-glutamine (Gibco #25030081), 0.5X B-27 supplement without Vitamin A (Gibco #12587010), and 0.5X N-2 supplement (Gibco #17502048)) supplemented with 0.05 mg/ml ascorbic acid (Sigma-Aldrich #A4544) and 4.53×10^{-4} M monothioglycerol (Sigma-Aldrich #M6145). EBs were broken up by trypsinization at 48 hours and re-plated at the same density for an additional 40 hours, with supplementation of 5 ng/ml VEGF (R&D Systems #293-VE), 8 ng/ml activin A (R&D Systems #338-AC), and 0.1 ng/ml BMP-4 (R&D Systems #314-BP) in addition to ascorbic acid and monothioglycerol as before, after which pre-cardiac mesoderm (MES) was collected. EBs were then broken up by trypsinization and plated in a monolayer on tissue culture-treated, gelatinized plates at 450,000 cells/cm^2 in StemPro-34 (Gibco #10639011) supplemented with 5 ng/mL VEGF, 10 ng/mL human FGF basic (R&D Systems #233-FB), 25 ng/mL FGF-10 (R&D Systems #345-FG), and 0.1 mg/ml ascorbic acid. Cardiac progenitor (CP) cells were collected 32 hours after plating. Media was then changed twice a day for the next two days, and once a day subsequently, until collection of CMs 10 days after the start. To suppress growth of non-cardiomyocyte cells, cytokine and growth factor supplementation was stopped at the end of day 6 of the differentiation protocol.

Success of differentiation was evaluated by staining for particular markers at each stage, looking for purity thresholds by FACS analysis. At the MES stage, double-staining was performed for Flk-1/Kdr (anti-mouse FLK-1, APC-conjugated, BD Biosciences #560070, 1:50 dilution) and Pdgfr-α (anti-mouse CD140a, PE-conjugated, Thermo Fisher #12-1401-81, 1:100 dilution), with a typical successful differentiation containing at least 50% double-positive cells.

At the CP stage, FACS analysis for GFP was performed, with a typical successful differentiation containing at least 60% GFP-positive cells. At the CM stage, staining for cTnT (primary: anti-cardiac troponin T, Abcam #ab8295, 1:200; secondary: rabbit anti-mouse IgG, Alexa Fluor 647-conjugated, Invitrogen #A21239, 1:100) was performed, with a typical successful differentiation containing at least 70% positive cells. FACS analysis was performed on a BD LSR II HTS analyzer.

ATAC-seq

The initial set of ATAC-seq experiments were performed according to a protocol based on the original ATAC-seq protocol (Buenrostro et al., 2015). These samples were sequenced by paired-end with 40bp reads on an Illumina HiSeq 2000.

The latter set of ATAC-seq experiments were performed according to a protocol based on the OmniATAC protocol (Corces et al., 2017). Briefly, cells were trypsinized from culture and washed twice in cold PBS, then 100,000 cells were resuspended in 100 ul cold lysis buffer (10mM Tris-HCl, pH 7.5, 10mM NaCl, 3mM $MgCl_2$, 0.1% NP-40, 0.1% Tween-20, 0.01% digitonin). Cells were then incubated on ice for 3 minutes, then 1 ml of cold wash buffer (10mM Tris-HCl, pH 7.5, 10mM NaCl, 3mM $MgCl_2$, 0.1% Tween-20) was added. Cells were then centrifuged at 4°C for 10 minutes at 500 x g, and the supernatant was carefully discarded. Next, cells were thoroughly resuspended by pipetting in 50 ul transposition reaction mix (1x Tagment DNA Buffer (Illumina #15027866), 0.33X PBS, 0.1% Tween-20, 0.01% digitonin, 5ul Tagment DNA Enzyme 1 (Illumina #20034197)). Reactions were incubated at 37°C with shaking at 300 rpm for 30 minutes, then DNA was immediately isolated using the Qiagen MinElute Reaction Cleanup Kit (Qiagen #28204) and eluted in 10 ul of the elution buffer. Libraries were prepared as previously described (Buenrostro et al., 2015), with the exception of the qPCR side reaction to determine optimal final amplification, which uses 0.83 uM of each primer (Ad1_noMX and the Ad2.* indexing primer) and 1X SYBR Green (Thermo Fisher #S7563), and the final purification and size selection, which was performed using Agencourt AMPure XP beads (Beckman Coulter #A63880). Library quality control was performed prior to sequencing on an Advanced Analytical Fragment Analyzer, and concentration was determined by qPCR on a

LightCycler 480 (Roche). These samples were sequenced by paired-end with 75bp reads on an Illumina NextSeq 500.

ATAC-seq data processing

ATAC-seq experiments were processed through the use of the ENCODE DCC ATAC-seq processing pipeline (Dunham et al., 2012), available at https://github.com/ENCODE-DCC/atac-seq-pipeline, starting with BAM files aligned to the NCBI Build 37 (mm9) reference genome using the Burrows-Wheeler aligner (Li and Durbin, 2009). High-confidence peaks were called using the irreproducible discovery rate (IDR) method, with a cutoff of 0.05, in each stage. Overlapping peaks were then collapsed and merged using the mergeBed function in the BEDtools suite (Quinlan and Hall, 2010). Peak summits were identified on peaks present in each stage using the MACS2 refinepeak subcommand (Zhang et al., 2008). ATAC-seq data from fetal ventricles from Akerberg et al. (2019) and adult ventrical apices from Quaife-Ryan et al. (2017) were run through the pipeline for display purposes and correlation calculations; these data are available at the Gene Expression Omnibus (GEO) under accession numbers GSE124008 and GSE95764, respectively.

ATAC-seq analysis

Correlations between replicates were calculated over the union set of ATAC-seq peaks across all four stages of *in vitro* CM differentiation, using Spearman's rho with the plotCorrelation program in the deepTools suite on the input BAM files (Ramírez et al., 2016). All other correlations were also performed using deepTools on bigWig files produced by the processing pipeline, over an appropriate set of regions as specified.

Identification of differentially enriched open chromatin regions was performed using the getDifferentialPeaks function from the HOMER suite (Heinz et al., 2010). ATAC-seq peaks open in CM that intersect with active enhancers were used as the input regions, with the fold-change cut-off of 1.5 (-F 1.5) between DMSO-treated and dTAG-treated CMs. All other parameters were kept at default settings.

Boxplots were generated using the mean ATAC-seq signal over each TEAD1 peak in an active enhancer, with the output from the multiBigwigSummary function of the deepTools suite.

Identification of putative *cis*-regulatory elements

Enhancers were annotated using the ChromHMM algorithm (Ernst et al., 2011). Briefly, ChIP-seq data from all 4 stages (ESC, MES, CP, and CM) for the histone modifications H3K27me3, H3K4me1, H3K27ac, H3K4me3, as well as RNA polymerase II, phosphorylated at serine 5 (RNAPSer5), and pooled whole cell extracts as inputs (Wamstad et al., 2012) with default parameters, except for the number of states. We tested a range of states, choosing 11 as it provided optimal discrimination between promoter states and enhancer states. Chromatin states were then annotated based on enrichment for characteristic marks, using some guidance from the original ChromHMM study (Ernst et al., 2011).

Super-enhancers were called using the ROSE algorithm (Lovén et al., 2013; Whyte et al., 2013). H3K27ac peaks were called in all four stages of CM differentiation using the ENCODE DCC ChIP-seq pipeline with an IDR cutoff of 0.05, using our previously-generated H3K27ac ChIP-seq data (Wamstad et al., 2012). These peaks were then used as input into the stitching algorithm, with default parameters (2.5 kb TSS exclusion distance, 12.5 kb stitching distance).

Identification of enriched transcription factor motifs

The findMotifsGenome command, from the HOMER suite (Heinz et al., 2010) was used to identify enriched TF motifs, with using the repeat-masked version of the mm9 genome (-mask) and a 500bp region around peak summits (-size 500). All other parameters were used as default settings. De novo motifs are reported, along with their best match from JASPAR according to HOMER's alignments, prioritizing strong matches with the core of the motif and stage-specific TF gene expression as appropriate.

Generation of transgenic ESC lines

Mouse embryonic stem cells (E14 Tg Nkx2-5-EmGFP) (Hsiao et al., 2008) cultured on feeders

using standard technique, were transfected using Lipofectamine 3000 (Thermo Fisher #L3000001), in the absence of feeders. Transfection reactions used 5 ug of material, with a 1:1 ratio between donor plasmid and guide/Cas9-GFP plasmid (pX458), prepared for each specific guide according to instructions (found at https://www.addgene.org/crispr/zhang/) using oligonucleotides indicated below (Ran et al., 2013). 24 hours after transfection, cells were bulk sorted by FACS for GFP expression and plated on feeders. Once these cells expanded sufficiently, after 4-5 days, they were sorted into single cells by the included fluorescent marker. Single colonies were then expanded and genotyed for proper insertion by PCR and sequence verified.

Conditional suppression of TEAD1 protein levels

100 nM of dTAG-13 (Tocris Biosciences #6605) was added to the cell culture media. An equivalent amount of DMSO was added to control samples. Depletion of TEAD1 was measured by Western blot with antibodies to HA (Cell Signaling #3724S) on nuclear lysates isolated by standard procedure (Wysocka et al., 2001).

Chromatin immunoprecipitation

ChIP was performed in about 10 million cells per replicate, as previously described (Mylonas and Tessarz, 2018), using either an anti-HA antibody (Cell Signaling #3724) or an anti-H3K27ac antibody (Abcam #ab4729). Briefly, cells were cross-linked with 1% formaldehyde for 10 minutes followed by 5 minutes quenching with 125 mM glycine. After washing with PBS buffer, the cells were collected and lysed in cell lysis buffer (5 mM Tris pH 8.0, 85 mM KCl, and 0.5% NP-40) with x1 Halt Protease Inhibitor cocktail (ThermoFisher 87786) and 1mM PMSF (Sigma 10837091001). Pellets were spun for 5 minutes at 6000 rpm at 4°C. Nuclei were lysed in nuclei lysis buffer (1% SDS, 10 mM EDTA, and 50 mM Tris–HCl) and samples were sonicated for 12 min on a Covaris Sonicator. The samples were centrifuged for 20 minutes at 13,000 rpm at 4°C and the supernatant was diluted in IP buffer (0.01% SDS, 1.1% Triton-X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl, and 167 mM NaCl), and the appropriate antibody (10 μg) was added and incubated overnight at 4°C with rotation. Two biological replicates were prepared for each condition using

independent cell cultures and chromatin precipitations. The next day, 50 µl Protein G Dynabeads (Life Technologies #10009D) were added for 1 hour at room temperature with rotation. Beads were washed once for 1 min with rotation with each of the following buffers: Low salt buffer (0.1% SDS, 1% Triton-X-100, 2mM EDTA, 20mM Tris pH 8.0, 150mM NaCl), High salt buffer (0.1% SDS, 1% Triton-X-100, 2mM EDTA, 20mM Tris pH 8.0, 500mM NaCl), LiCl buffer (0.25M LiCL, 1% NP-40, 1% deoxycholate, 10mM Tris pH 8.0, 1mM EDTA), and TE buffer (50mM Tris pH 8.0, 10mM EDTA). DNA was eluted off the beads by rotation at room temperature for 30 min in 200 µL elution buffer (1% SDS, 0.1M NaHCO3). Cross-links were reversed at 65°C for 4h. RNA was degraded by the addition of 3 µL of 33 mg/mL RNase A (Sigma, R4642) and incubation at 37°C for 2 hours. Protein was degraded by the addition of 5 µL of 20 mg/mL proteinase K (Invitrogen #25530049) and incubation at 50°C for 1 hour. A phenol:chloroform:isoamyl alcohol extraction was performed followed by an ethanol precipitation. The DNA was then resuspended in 20 µL water and used for either qPCR or sequencing.

For ChIP-seq experiments, purified 10-20 ng of ChIP DNA was used to prepare Illumina multiplexed sequencing libraries. Libraries for Illumina sequencing were prepared following the NEBNext DNA Library Prep Master Mix kit (NEB E6040). Amplified libraries were size-selected using a 2% agarose gel to extract fragments between 200 and 600 bp.

ChIP-seq data processing

ChIP-seq experiments were processed through the use of the ENCODE DCC ChIP-seq processing pipeline, available at https://github.com/ENCODE-DCC/chip-seq-pipeline2. Inputs to the pipeline were BAM files aligned to the NCBI Build 37 (mm9) reference genome, using the Burrows-Wheeler aligner (Li and Durbin, 2009). High-confidence peaks were called using overlaps between replicates for histone experiments. Data from Wamstad et al. (2012) used in this study is available at https://b2b.hci.utah.edu/gnomex/ using a guest login, or at GEO under the accession number GSE47949. For the binding of cardiac TFs except HAND2, peak summits from Akerberg et al. (2019) were used, and raw sequencing data (available under the accession number GSE124008 from GEO) was run through the pipeline for display purposes. For HAND2

binding, raw sequencing data from Laurent et al. (2017) (available from GEO at accession number GSE73368) was run through the pipeline and peaks were called using the IDR approach described for ATAC-seq peaks above.

ChIP-seq analysis

Identification of differentially enriched H3K27ac regions was performed using the DiffBind R package (Ross-Innes et al., 2012). Input regions surveyed were peaks from prior CM H3K27ac ChIP-seq experiments (Wamstad et al., 2012) that overlap with ChromHMM-annotated active enhancers. Differential counts were determined using the edgeR-based method included in the package, and significantly changed regions were called based on a 10% FDR.

Boxplots were generated using the mean H3K27ac signal over each active enhancer bound by TEAD1, with the output from the multiBigwigSummary function of the deepTools suite.

Antibodies

Antibody	Company	Identifier
Mouse anti-histone H3	Abcam	ab1791
Goat anti-Lamin B	Santa Cruz Biotechnology	sc-6217
Rabbit anti-HA	Cell Signaling	3724S
Rabbit anti-H3K27ac	Abcam	ab4729
Mouse anti-TEAD1	BD Biosciences	610922
Mouse anti-cTnT	Abcam	ab8295
Rabbit anti-GAPDH	Santa Cruz Biotechnology	sc-25778

Oligonucleotides

Identification	Sequence
TEAD1 sg1 F	CACCGatctctctatataaataacc
TEAD1 sg1 R	AAACggttatttatatagagagatC
TEAD1 sg2 F	CACCGcatatctacaggcttgtga
TEAD1 sg2 R	AAACtcacaagcctgtagatatgC
TEAD1 sg3 F	CACCGccttcacaagcctgtagata

TEAD1 sg3 R	AAACtatctacaggcttgtgaaggC
TEAD1 sg4 F	CACCgagcacagcaccatatctac
TEAD1 sg4 R	AAACgtagatatggtgctgtgctc
TEAD1-FKBP genotype primer F	AATGTGGCACAGATTGACCA
TEAD1-FKBP genotype primer R	ATTCACCTTCCCCAAGGAAT

REFERENCES

Adachi, K., Kopp, W., Wu, G., Heising, S., Greber, B., Stehling, M., Araúzo-Bravo, M.J., Boerno, S.T., Timmermann, B., Vingron, M., et al. (2018). Esrrb Unlocks Silenced Enhancers for Reprogramming to Naive Pluripotency. Cell Stem Cell *23*, 266-275.e6.

Afouda, B.A., Martin, J., Liu, F., Ciau-Uitz, A., Patient, R., and Hoppler, S. (2008). GATA transcription factors integrate Wnt signalling during heart development. Development *135*, 3185–3190.

Ahmad, S.M., Bhattacharyya, P., Jeffries, N., Gisselbrecht, S.S., and Michelson, A.M. (2015). Two Forkhead transcription factors regulate cardiac progenitor specification by controlling the expression of receptors of the fibroblast growth factor and Wnt signaling pathways. Dev Camb Engl *143*, 306–317.

Akerberg, B.N., Gu, F., VanDusen, N.J., Zhang, X., Dong, R., Li, K., Zhang, B., Zhou, B., Sethi, I., Ma, Q., et al. (2019). A reference map of murine cardiac transcription factor chromatin occupancy identifies dynamic and conserved enhancers. Nat Commun *10*, 4907.

Azakie, A., LaMont, L., Fineman, J.R., and He, Y. (2005). Divergent transcriptional enhancer factor-1 regulates the cardiac troponin T promoter. American Journal of Physiology-Cell Physiology *289*, C1522–C1534.

Blue, G.M., Kirk, E.P., Giannoulatou, E., Sholler, G.F., Dunwoodie, S.L., Harvey, R.P., and Winlaw, D.S. (2017). Advances in the Genetics of Congenital Heart Disease: A Clinician's Guide. J Am Coll Cardiol *69*, 859–870.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. Cell *122*, 947–956.

Bruneau, B.G. (2013). Signaling and Transcriptional Networks in Heart Development and Regeneration. Cold Spring Harbor Perspectives in Biology *5*, a008292–a008292.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature Methods *10*, 1213–1218.

Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide (John Wiley & Sons, Inc.).

Chambers, A.E., Logan, M., Kotecha, S., Towers, N., Sparrow, D., and Mohun, T.J. (1994). The RSRF/MEF2 protein SL1 regulates cardiac muscle-specific transcription of a myosin light-chain gene in Xenopus embryos. Genes & Development *8*, 1324–1334.

Chang, L., Azzolin, L., Biagio, D.D., Zanconato, F., Battilana, G., Xiccato, R.L., Aragona, M., Giulitti, S., Panciera, T., Gandin, A., et al. (2018). The SWI/SNF complex is a mechanoregulated inhibitor of YAP and TAZ. Nature *563*, 265–269.

Chen, H.-H., Mullett, S.J., and Stewart, A.F.R. (2004a). Vgl-4, a Novel Member of the Vestigial-like Family of Transcription Cofactors, Regulates α1-Adrenergic Activation of Gene Expression in Cardiac Myocytes. J Biol Chem *279*, 30800–30806.

Chen, H.-H., Baty, C.J., Maeda, T., Brooks, S., Baker, L.C., Ueyama, T., Gursoy, E., Saba, S., Salama, G., London, B., et al. (2004b). Transcription Enhancer Factor-1-Related Factor-Transgenic Mice Develop Cardiac Conduction Defects Associated With Altered Connexin Phosphorylation. Circulation *110*, 2980–2987.

Chen, Z., Friedrich, G.A., and Soriano, P. (1994). Transcriptional enhancer factor 1 disruption by a retroviral gene trap leads to heart defects and embryonic lethality in mice. Gene Dev *8*, 2293–2301.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nature Methods *14*, 959–962.

Costello, I., Pimeisl, I.-M., Dräger, S., Bikoff, E.K., Robertson, E.J., and Arnold, S.J. (2011). The T-box transcription factor Eomesodermin acts upstream of Mesp1 to specify cardiac mesoderm during mouse gastrulation. Nature Cell Biology *13*, 1084–1091.

Costello, I., Nowotschin, S., Sun, X., Mould, A.W., Hadjantonakis, A.-K., Bikoff, E.K., and Robertson, E.J. (2015). Lhx1 functions together with Otx2, Foxa2, and Ldb1 to govern anterior mesendoderm, node, and midline development. Genes & Development *29*, 2108–2122.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences *107*, 21931–21936.

Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J., Kaul, R., et al. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

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

Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoeppner, D.J., Dash, C., Bazett-Jones, D.P., Grice, S.L., McKay, R.D.G., Buetow, K.H., et al. (2008). Global Transcription in Pluripotent

Embryonic Stem Cells. Cell Stem Cell *2*, 437–447.

Ege, N., Dowbaj, A.M., Jiang, M., Howell, M., Hooper, S., Foster, C., Jenkins, R.P., and Sahai, E. (2018). Quantitative Analysis Reveals that Actin and Src-Family Kinases Regulate Nuclear YAP1 and Its Export. Cell Syst *6*, 692-708.e13.

Elosegui-Artola, A., Andreu, I., Beedle, A.E.M., Lezamiz, A., Uroz, M., Kosmalska, A.J., Oria, R., Kechagia, J.Z., Rico-Lastres, P., Roux, A.-L.L., et al. (2017). Force Triggers YAP Nuclear Entry by Regulating Transport across Nuclear Pores. Cell *171*, 1397-1410.e14.

Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., et al. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. Nature *473*, 43–49.

Felsenfeld, G., Boyes, J., Chung, J., Clark, D., and Studitsky, V. (1996). Chromatin structure and gene expression. Proceedings of the National Academy of Sciences *93*, 9384–9388.

Festuccia, N., Owens, N., and Navarro, P. (2017). Esrrb, an estrogen-related receptor involved in early development, pluripotency and reprogramming. Febs Lett *592*, 852–877.

Furlong, E.E.M., and Levine, M. (2018). Developmental enhancers and chromosome topology. Science *361*, 1341–1345.

Gelb, B.D., and Chung, W.K. (2014). Complex Genetics and the Etiology of Human Congenital Heart Disease. Cold Spring Harbor Perspectives in Medicine *4*, a013953–a013953.

Gerstein, M.B., Kundaje, A., Hariharan, M., Landt, S.G., Yan, K.-K., Cheng, C., Mu, X.J., Khurana, E., Rozowsky, J., Alexander, R., et al. (2012). Architecture of the human regulatory network derived from ENCODE data. Nature *489*, 91–100.

Gotea, V., Visel, A., Westlund, J.M., Nobrega, M.A., Pennacchio, L.A., and Ovcharenko, I. (2010). Homotypic clusters of transcription factor binding sites are a key component of human promoters and enhancers. Genome Research *20*, 565–577.

Grigoryev, S.A., Bulynko, Y.A., and Popova, E.Y. (2006). The end adjusts the means: Heterochromatin remodelling during terminal cell differentiation. Chromosome Res *14*, 53–69.

Gyoja, F. (2017). Basic helix-loop-helix transcription factors in evolution: Roles in development of mesoderm and neural tissues. Genesis *55*, e23051.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Molecular Cell *38*, 576–589.

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-Enhancers in the Control of Cell Identity and Disease. Cell *155*, 934–947.

Homsy, J., Zaidi, S., Shen, Y., Ware, J.S., Samocha, K.E., Karczewski, K.J., DePalma, S.R., McKean, D., Wakimoto, H., Gorham, J., et al. (2015). De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. Sci New York N Y *350*, 1262–1266.

Hsiao, E.C., Yoshinaga, Y., Nguyen, T.D., Musone, S.L., Kim, J.E., Swinton, P., Espineda, I., Manalac, C., deJong, P.J., and Conklin, B.R. (2008). Marking Embryonic Stem Cells with a 2A Self-Cleaving Peptide: A NKX2-5 Emerald GFP BAC Reporter. Plos One *3*, e2532.

Jin, S.C., Homsy, J., Zaidi, S., Lu, Q., Morton, S., DePalma, S.R., Zeng, X., Qi, H., Chang, W., Sierant, M.C., et al. (2017). Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. Nat Genet *49*, 1593–1601.

Kaneko, K.J., Kohn, M.J., Liu, C., and DePamphilis, M.L. (2007). Transcription factor TEAD2 is involved in neural tube closure. Genesis *45*, 577–587.

Kang, J., Nathan, E., Xu, S.-M., Tzahor, E., and Black, B.L. (2009). Isl1 is a direct transcriptional target of Forkhead transcription factors in second-heart-field-derived mesoderm. Dev Biol *334*, 513–522.

Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., and Weirauch, M.T. (2018). The Human Transcription Factors. Cell *172*, 650–665.

Laurent, F., Girdziusaite, A., Gamart, J., Barozzi, I., Osterwalder, M., Akiyama, J.A., Lincoln, J., Lopez-Rios, J., Visel, A., Zuniga, A., et al. (2017). HAND2 Target Gene Regulatory Networks Control Atrioventricular Canal and Cardiac Valve Development. Cell Reports *19*, 1602–1613.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinform Oxf Engl *25*, 1754–1760.

Lovén, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers. Cell *153*, 320–334.

Luna-Zurita, L., Stirnimann, C.U., Glatt, S., Kaynak, B.L., Thomas, S., Baudin, F., Samee, M.A.H., He, D., Small, E.M., Mileikovsky, M., et al. (2016). Complex Interdependence Regulates Heterotypic Transcription Factor Distribution and Coordinates Cardiogenesis. Cell *164*, 999–1014.

Maeda, T., Mazzulli, J.R., Farrance, I.K.G., and Stewart, A.F.R. (2002). Mouse DTEF-1 (ETFR-1, TEF-5) Is a Transcriptional Activator in α 1 -Adrenergic Agonist-stimulated Cardiac Myocytes. J

Biol Chem *277*, 24346–24352.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nature Biotechnology *28*, 495–501.

Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T., and Misteli, T. (2006). Hyperdynamic Plasticity of Chromatin Proteins in Pluripotent Embryonic Stem Cells. Dev Cell *10*, 105–116.

Mosqueira, D., Pagliari, S., Uto, K., Ebara, M., Romanazzo, S., Escobedo-Lucea, C., Nakanishi, J., Taniguchi, A., Franzese, O., Nardo, P.D., et al. (2014). Hippo Pathway Effectors Control Cardiac Progenitor Cell Fate by Acting as Dynamic Sensors of Substrate Mechanics and Nanostructure. ACS Nano *8*, 2033–2047.

Mylonas, C., and Tessarz, P. (2018). Transcriptional repression by FACT is linked to regulation of chromatin accessibility at the promoter of ES cells. Life Sci Alliance *1*, e201800085.

Nabet, B., Roberts, J.M., Buckley, D.L., Paulk, J., Dastjerdi, S., Yang, A., Leggett, A.L., Erb, M.A., Lawlor, M.A., Souza, A., et al. (2018). The dTAG system for immediate and target-specific protein degradation. Nat Chem Biol *14*, 431–441.

Natarajan, A., Yardimci, G.G., Sheffield, N.C., Crawford, G.E., and Ohler, U. (2012). Predicting cell-type-specific gene expression from regions of open chromatin. Genome Research *22*, 1711–1722.

Neph, S., Vierstra, J., Stergachis, A.B., Reynolds, A.P., Haugen, E., Vernot, B., Thurman, R.E., John, S., Sandstrom, R., Johnson, A.K., et al. (2012). An expansive human regulatory lexicon encoded in transcription factor footprints. Nature *489*, 83–90.

Nishioka, N., Yamamoto, S., Kiyonari, H., Sato, H., Sawada, A., Ota, M., Nakao, K., and Sasaki, H. (2008). Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. Mechanisms of Development *125*, 270–283.

Nord, A.S., Blow, M.J., Attanasio, C., Akiyama, J.A., Holt, A., Hosseini, R., Phouanenavong, S., Plajzer-Frick, I., Shoukry, M., Afzal, V., et al. (2013). Rapid and Pervasive Changes in Genome-wide Enhancer Usage during Mammalian Development. Cell *155*, 1521–1531.

Ota, M., and Sasaki, H. (2008). Mammalian Tead proteins regulate cell proliferation and contact inhibition as transcriptional mediators of Hippo signaling. Development *135*, 4059–4069.

Paige, S.L., Thomas, S., Stoick-Cooper, C.L., Wang, H., Maves, L., Sandstrom, R., Pabon, L., Reinecke, H., Pratt, G., Keller, G., et al. (2012). A Temporal Chromatin Signature in Human Embryonic Stem Cells Identifies Regulators of Cardiac Development. Cell *151*, 221–232.

Peterkin, T. (2003). GATA-6 maintains BMP-4 and Nkx2 expression during cardiomyocyte precursor maturation. The EMBO Journal *22*, 4260–4273.

Postma, A.V., Bezzina, C.R., and Christoffels, V.M. (2016). Genetics of congenital heart disease: the contribution of the noncoding regulatory genome. Journal of Human Genetics *61*, 13–19.

Potthoff, M.J., and Olson, E.N. (2007). MEF2: a central regulator of diverse developmental programs. Development *134*, 4131–4140.

Prendiville, T., Jay, P.Y., and Pu, W.T. (2014). Insights into the Genetic Structure of Congenital Heart Disease from Human and Murine Studies on Monogenic Disorders. Cold Spring Harbor Perspectives in Medicine *4*, a013946–a013946.

Purvis, J.E., and Lahav, G. (2013). Encoding and Decoding Cellular Information through Signaling Dynamics. Cell *152*, 945–956.

Quaife-Ryan, G.A., Sim, C.B., Ziemann, M., Kaspi, A., Rafehi, H., Ramialison, M., El-Osta, A., Hudson, J.E., and Porrello, E.R. (2017). Multicellular Transcriptional Analysis of Mammalian Heart Regeneration. Circulation *136*, 1123–1139.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinform Oxf Engl *26*, 841–842.

Rada-Iglesias, A., Bajpai, R., Prescott, S., Brugmann, S.A., Swigut, T., and Wysocka, J. (2012). Epigenomic Annotation of Enhancers Predicts Transcriptional Regulators of Human Neural Crest. Cell Stem Cell *11*, 633–648.

Ragni, C.V., Diguet, N., Garrec, J.-F.L., Novotova, M., Resende, T.P., Pop, S., Charon, N., Guillemot, L., Kitasato, L., Badouel, C., et al. (2017). Amotl1 mediates sequestration of the Hippo effector Yap1 downstream of Fat4 to restrict heart growth. Nat Commun *8*, 14582.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res *44*, W160-5.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc *8*, 2281–2308.

Reiter, F., Wienerroither, S., and Stark, A. (2017). Combinatorial function of transcription factors and cofactors. Curr Opin Genet Dev *43*, 73–81.

Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., et al. (2012). Differential oestrogen receptor binding is

associated with clinical outcome in breast cancer. Nature *481*, 389–393.

Sawada, A., Kiyonari, H., Ukita, K., Nishioka, N., Imuta, Y., and Sasaki, H. (2008). Redundant roles of Tead1 and Tead2 in notochord development and the regulation of cell proliferation and survival. Mol Cell Biol *28*, 3177–3189.

Scuderi, G.J., and Butcher, J. (2017). Naturally Engineered Maturation of Cardiomyocytes. Frontiers Cell Dev Biology *5*, 50.

Smemo, S., Campos, L.C., Moskowitz, I.P., Krieger, J.E., Pereira, A.C., and Nobrega, M.A. (2012). Regulatory variation in a TBX5 enhancer leads to isolated congenital heart disease. Human Molecular Genetics *21*, 3255–3263.

Sonawane, A.R., Platig, J., Fagny, M., Chen, C.-Y., Paulson, J.N., Lopes-Ramos, C.M., DeMeo, D.L., Quackenbush, J., Glass, K., and Kuijjer, M.L. (2017). Understanding Tissue-Specific Gene Regulation. Cell Reports *21*, 1077–1088.

Spitz, F., and Furlong, E.E.M. (2012). Transcription factors: from enhancer binding to developmental control. Nature Reviews Genetics *13*, 613–626.

Stergachis, A.B., Neph, S., Reynolds, A., Humbert, R., Miller, B., Paige, S.L., Vernot, B., Cheng, J.B., Thurman, R.E., Sandstrom, R., et al. (2013). Developmental Fate and Cellular Maturity Encoded in Human Regulatory DNA Landscapes. Cell *154*, 888–903.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell *126*, 663–676.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. Nature *489*, 75–82.

Uosaki, H., Cahan, P., Lee, D.I., Wang, S., Miyamoto, M., Fernandez, L., Kass, D.A., and Kwon, C. (2015). Transcriptional Landscape of Cardiomyocyte Maturation. Cell Reports *13*, 1705–1716.

Vermij, S.H., Abriel, H., and Veen, T.A.B. van (2017). Refining the molecular organization of the cardiac intercalated disc. Cardiovasc Res *113*, cvw259.

Vierbuchen, T., Ling, E., Cowley, C.J., Couch, C.H., Wang, X., Harmin, D.A., Roberts, C.W.M., and Greenberg, M.E. (2017). AP-1 Transcription Factors and the BAF Complex Mediate Signal-Dependent Enhancer Selection. Molecular Cell *68*, 1067-1082.e12.

Voss, T.C., and Hager, G.L. (2013). Dynamic regulation of transcriptional states by chromatin and transcription factors. Nat Rev Genetics *15*, 69–81.

Wamstad, J.A., Alexander, J.M., Truty, R.M., Shrikumar, A., Li, F., Eilertson, K.E., Ding, H., Wylie, J.N., Pico, A.R., Capra, J.A., et al. (2012). Dynamic and Coordinated Epigenetic Regulation of Developmental Transitions in the Cardiac Lineage. Cell *151*, 206–220.

Wang, X., Tucker, N.R., Rizki, G., Mills, R., Krijger, P.H., Wit, E. de, Subramanian, V., Bartell, E., Nguyen, X.-X., Ye, J., et al. (2016). Discovery and validation of sub-threshold genome-wide association study loci using epigenomic signatures. ELife *5*, 1061.

Weintraub, A.S., Li, C.H., Zamudio, A.V., Sigova, A.A., Hannett, N.M., Day, D.S., Abraham, B.J., Cohen, M.A., Nabet, B., Buckley, D.L., et al. (2017). YY1 Is a Structural Regulator of Enhancer-Promoter Loops. Cell *171*, 1573-1588.e28.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell *153*, 307–319.

Winata, C.L., Kondrychyn, I., Kumar, V., Srinivasan, K.G., Orlov, Y., Ravishankar, A., Prabhakar, S., Stanton, L.W., Korzh, V., and Mathavan, S. (2013). Genome Wide Analysis Reveals Zic3 Interaction with Distal Regulatory Elements of Stage Specific Developmental Genes in Zebrafish. Plos Genet *9*, e1003852.

Wysocka, J., Reilly, P.T., and Herr, W. (2001). Loss of HCF-1-Chromatin Association Precedes Temperature-Induced Growth Arrest of tsBN67 Cells. Molecular and Cellular Biology *21*, 3820–3829.

Zaidi, S., Choi, M., Wakimoto, H., Ma, L., Jiang, J., Overton, J.D., Romano-Adesman, A., Bjornson, R.D., Breitbart, R.E., Brown, K.K., et al. (2013). De novo mutations in histone-modifying genes in congenital heart disease. Nature *498*, 220–223.

Zanconato, F., Forcato, M., Battilana, G., Azzolin, L., Quaranta, E., Bodega, B., Rosato, A., Bicciato, S., Cordenonsi, M., and Piccolo, S. (2015). Genome-wide association between YAP/TAZ/TEAD and AP-1 at enhancers drives oncogenic growth. Nature Cell Biology *17*, 1218–1227.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhu, Y., Sun, L., Chen, Z., Whitaker, J.W., Wang, T., and Wang, W. (2013). Predicting enhancer transcription and activity from chromatin modifications. Nucleic Acids Res *41*, 10032–10043.

Chapter 3

Discussion

Author Contributions for Chapter 3:
This chapter was prepared by Olukunle Demuren.

Summary and Future Directions

This thesis focused on determining the mechanisms by which molecular mediators regulate enhancer activation and tissue-specific gene expression patterns cardiac lineage commitment. Although the work presented in this thesis focuses on cardiac development, these insights are generalizable to a wide variety of developmental systems. Chapter 1 details the current understanding of transcriptional regulation of developmental gene expression, centered on enhancer-mediated gene expression and cardiac lineage commitment. Further discussion on the concept of transcription factors serving as mediators of extracellular signaling cues by binding at *cis*-regulatory elements was presented with an emphasis on the Hippo-TEAD axis in heart development. The experimental work in Chapter 2 leveraged many of the concepts and technological advances described in the introductory chapter. We developed an integrated system for the unbiased identification of novel regulators of cardiac cell fate. By measuring chromatin accessibility and histone modification patterns in a dynamic model of cardiac lineage commitment, we identified TEAD1 as a critical regulator of a selected set of enhancers that appear to target genes with roles in cell-cell contacts and cardiac contraction, a role not previously assigned to this TF.

The data presented here represent substantial insights into the mechanisms of cardiac-specific gene regulation. In Chapter 2, we first demonstrated that dynamic chromatin accessibility within enhancers can be used to identify candidate TFs that regulate lineage-specific enhancer activation. We then employed a targeted perturbation strategy, exploiting new developments in protein engineering (Nabet et al., 2018), to directly test TEAD1's role at enhancers. Our ability to deplete TEAD1 during cell fate transitions from cardiac progenitors (CPs) to CMs allowed us to specifically test its direct role in a temporal manner. Although depletion of TEAD1 as cells transition to CP and CM did not affect differentiation per se, we showed that the resulting CMs displayed an abnormal beating phenotype. Notably, TEAD1 depletion significantly decreased the levels of chromatin accessibility and active histone modifications at a subset of enhancers, most of which were bound by TEAD1. Upon further analysis, we posited that this set of enhancers targets genes with roles in the formation and maintenance of cell-cell junctions as

well as interactions with the actin cytoskeleton, consistent with our phenotypic observations. Ongoing efforts aim to connect changes in gene expression to the loss of TEAD1 binding at enhancers. Moreover, we expect that by conducting phenotypic rescue experiments with candidate TEAD1 target genes, we will develop a more complete model of TEAD1-mediated gene regulation during cardiac lineage commitment. Nevertheless, these studies represent a substantial advance in our understanding of the key TFs that drive context-specific cardiac lineage commitment.

Enhancers as integrators of signaling cues

Extracellular and intracellular signaling cues are transmitted into transcriptional outcomes by TF binding at enhancers. These responses are particularly important during development as the inputs of multiple signaling pathways must be integrated in individual cells to drive proper cell fate decisions. Heart development is a powerful example of this process; as discussed in Chapter 1, the activity of multiple signaling pathways is required to drive the differentiation of cells into the different tissues that make up the adult heart (Noseda et al., 2011). The example of canonical Wnt signaling, which switches from promoting cardiac cell fates through induction of a specific TF after a discrete timepoint during development (Liu et al., 2007; Naito et al., 2006; Ueno et al., 2007), illustrates the complexity and precision required for proper heart development. This paradigm is also illustrated by the action of the Hippo signaling pathway, which influences organ size and cell proliferation through the transcriptional co-activators YAP and TAZ that bind to TEAD TFs at target loci in the nucleus (Ota and Sasaki, 2008). Recent studies have revealed that YAP and TAZ drive transcriptional responses to mechanotransduction through TEAD1 (Azzolin et al., 2014; Chang et al., 2018; Ege et al., 2018; Paramasivam et al., 2011; Ragni et al., 2017; Totaro et al., 2017), so it will be important to dissect the set of target genes downstream of Hippo signaling and mechanotransduction.

Although at the writing of this thesis, we were unable to deconvolve the contribution of each of these pathways to TEAD1-mediated regulation of enhancer activity, our data suggest that this latter pathway may be highly relevant for later stages of cardiac development, given that we

did not observe significant changes in cell proliferation upon TEAD1 depletion (data not shown). Moreover, genes in proximity to enhancers most affected by TEAD1 depletion have roles in mechanical cell-cell junctions and the actin cytoskeleton, suggesting that cross-talk with mechanotransduction might be a key part of TEAD1's regulation of the developing heart. Although it is widely accepted that enhancer elements play vital roles in regulating spatiotemporal gene expression patterns in a context-dependent manner through the binding of sets of TFs in response to signaling cues, understanding how specific upstream signaling pathways target TFs to a selected set of enhancers remains an important, unresolved question.

Connecting transcription factor action at enhancers to developmental gene expression

Developmental enhancers are critical nodes in a broader regulatory network that must function properly in time and space, responding to all manner of conflicting stimuli to control the expression of cell-type-specific genes. It is thought that the sets of TFs bound at enhancers are critical for mediating this specificity, but we know very little about how TFs interact to govern binding at some sites and not at others, even when the same primary sequence motif is present. As discussed in Chapter 1, chromatin accessibility is a major determinant of whether TFs are able to bind at regulatory elements, thus serving as an indicator of TF occupancy. With our growing knowledge of TF-mediated enhancer regulation, the focus on the complexity of TF binding has deepened, identifying important determinants beyond primary sequence such as the shape of the DNA and the presence of co-factors. For example, the TEAD family contains multiple members which are thought to bind the same motif, as do many TF families with roles in heart development, such as MEF2 and GATA. Notably, this multiplicity seems to be a feature of higher metazoans as both MEF2 and TEAD TFs have only one identified homolog in *Drosophila* (Bour et al., 1995; Wu et al., 2008). It remains an open question how and why these different TFs that bind the same or similar primary sequences discriminate their sites of regulatory action when expressed in the same cell type. Further biochemical and biophysical studies coupled with bioinformatic methods to identify candidate TF partners to characterize these partnerships that drive specific gene expression programs are needed to get to the heart of this question.

The role of particular co-factors in controlling tissue-specific gene regulation must also be explored. Throughout this thesis, for simplicity's sake YAP1 and its paralog TAZ (encoded by the gene *Wwtr1*) have mostly been considered together with regards to their potential partnership with TEAD1 during cardiac lineage commitment. However, *Yap1*- and *Wwtr1*-null mice have distinct phenotypes (Hossain et al., 2007; Makita et al., 2008; Morin-Kensicki et al., 2005) suggesting that these co-activators are regulating distinct gene expression programs. Assessments of genome-wide binding, as well as targeted perturbation such as the strategy used in this study, can help determine their individual roles in transmitting signaling cues into transcriptional outcomes and whether these co-activators also have TEAD-independent functions.

Linking enhancer elements to target genes adds another layer of complexity to that has yet to be resolved in the field. Technologies designed to measured enhancer-gene linkages, whether through spatial proximity-based approaches like Hi-C or through CRISPR-based perturbation strategies, have consistently shown that lineage-specific genes are regulated by multiple enhancers (Fulco et al., 2016; Jin et al., 2013; Klann et al., 2017). These studies also show that at least for the specific genes tested, disrupting any one enhancer of a set does not completely abrogate the expression of a gene and the contributions of each enhancer vary in a context-dependent manner. TFs are known to show cooperativity at enhancers, and it follows from those examples that enhancer elements themselves may cooperate to regulate the expression of particular genes. For example, super-enhancer clusters might represent a distinct regulatory paradigm for cell-type-specific gene expression through a force-multiplying effect of having many TFs and transcriptional co-activators present at one location in in the genome. Given that these clusters appear to regulate essential cell-identity TFs, reinforcing the expression of these genes through positive feedback auto-regulatory loops can ensure robust maintenance of expression. More sophisticated analyses, in which multiple factors are manipulated at once, will be needed to determine the order of events and most critical factors for gene activation in each cell type.

The recent development of prime editing, which allows high-efficiency and low-error rate genome editing (Anzalone et al., 2019), may allow more detailed studies aimed at dissecting the contribution of specific motifs and motif syntax at enhancers. Recent developments in single-cell approaches that can measure changes in chromatin states and gene expression will also be crucial to these explorations. The ability to combine these genome-wide approaches should also be instrumental in dissecting how variants identified in GWA studies affect gene expression and ultimately phenotypic variation. The majority of single nucleotide polymorphisms (SNPs) identified in GWA studies are in non-coding regions (Nishizaki and Boyle, 2017). Although many of these variants alter TF binding through mutations in a motif (Kasowski et al., 2010), a large proportion do not, while still impacting TF binding (Deplancke et al., 2016), indicating that further studies are needed. In particular, understanding how TFs regulate specific genes during development and disease and connecting this knowledge with information on genetic variation may elucidate additional genetic culprits behind CHD and other developmental disorders.

In vitro differentiation as a model for *in vivo* development

In vitro ESC-based differentiation systems, such as the CM differentiation used in this thesis, can be very helpful in dissecting temporal events that would otherwise be difficult to resolve *in vivo*. However, the fact that the heart and most organs must undergo substantial three-dimensional morphological changes during development that cannot be observed in a standard cell culture model presents challenges in extrapolating results to *in vivo* events. Nevertheless, *in vitro* studies can help identify molecular mechanisms and prioritize factors with potential roles in development and disease that can be tested *in vivo*. As discussed in Chapter 1, many of the genes currently linked to CHDs are TFs or chromatin modifiers (Homsy et al., 2015; Jin et al., 2017; Zaidi et al., 2013), although most of the genetic causes of CHD remain unknown (Blue et al., 2017). Thus, establishing the blueprint detailing the regulatory networks that connect specific TFs to genes and tissues within the heart is critical for interpreting data from genome-wide approaches as well as the potential of genetic variants as culprits of CHD. Single-cell studies will prove very useful in this regard and will allow *in vivo* analyses at a scale that was not

previously possible. In support of this idea, the ability to collect cells at discrete time points within the developing mouse heart enabled the reconstruction of key developmental events from gene expression data in a native context (Soysa et al., 2019). We expect that the studies and approaches discussed here, when combined with human genetic data from GWA studies linked to a variety of traits and diseases, will deepen our understanding of human development, the etiology of disease, and evolution.

References

Anzalone, A.V., Randolph, P.B., Davis, J.R., Sousa, A.A., Koblan, L.W., Levy, J.M., Chen, P.J., Wilson, C., Newby, G.A., Raguram, A., et al. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. Nature *576*, 149–157.

Azzolin, L., Panciera, T., Soligo, S., Enzo, E., Bicciato, S., Dupont, S., Bresolin, S., Frasson, C., Basso, G., Guzzardo, V., et al. (2014). YAP/TAZ incorporation in the β-catenin destruction complex orchestrates the Wnt response. Cell *158*, 157–170.

Blue, G.M., Kirk, E.P., Giannoulatou, E., Sholler, G.F., Dunwoodie, S.L., Harvey, R.P., and Winlaw, D.S. (2017). Advances in the Genetics of Congenital Heart Disease: A Clinician's Guide. J Am Coll Cardiol *69*, 859–870.

Bour, B.A., O'Brien, M.A., Lockwood, W.L., Goldstein, E.S., Bodmer, R., Taghert, P.H., Abmayr, S.M., and Nguyen, H.T. (1995). Drosophila MEF2, a transcription factor that is essential for myogenesis. Gene Dev *9*, 730–741.

Chang, L., Azzolin, L., Biagio, D.D., Zanconato, F., Battilana, G., Xiccato, R.L., Aragona, M., Giulitti, S., Panciera, T., Gandin, A., et al. (2018). The SWI/SNF complex is a mechanoregulated inhibitor of YAP and TAZ. Nature *563*, 265–269.

Deplancke, B., Alpern, D., and Gardeux, V. (2016). The Genetics of Transcription Factor DNA Binding Variation. Cell *166*, 538–554.

Ege, N., Dowbaj, A.M., Jiang, M., Howell, M., Hooper, S., Foster, C., Jenkins, R.P., and Sahai, E. (2018). Quantitative Analysis Reveals that Actin and Src-Family Kinases Regulate Nuclear YAP1 and Its Export. Cell Syst *6*, 692-708.e13.

Fulco, C.P., Munschauer, M., Anyoha, R., Munson, G., Grossman, S.R., Perez, E.M., Kane, M., Cleary, B., Lander, E.S., and Engreitz, J.M. (2016). Systematic mapping of functional enhancer-promoter connections with CRISPR interference. Science *354*, 769–773.

Homsy, J., Zaidi, S., Shen, Y., Ware, J.S., Samocha, K.E., Karczewski, K.J., DePalma, S.R., McKean, D., Wakimoto, H., Gorham, J., et al. (2015). De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. Sci New York N Y *350*, 1262–1266.

Hossain, Z., Ali, S.M., Ko, H.L., Xu, J., Ng, C.P., Guo, K., Qi, Z., Ponniah, S., Hong, W., and Hunziker, W. (2007). Glomerulocystic kidney disease in mice with a targeted inactivation of Wwtr1. Proc National Acad Sci *104*, 1631–1636.

Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.-A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature *503*, 290–294.

Jin, S.C., Homsy, J., Zaidi, S., Lu, Q., Morton, S., DePalma, S.R., Zeng, X., Qi, H., Chang, W., Sierant, M.C., et al. (2017). Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. Nat Genet *49*, 1593–1601.

Kasowski, M., Grubert, F., Heffelfinger, C., Hariharan, M., Asabere, A., Waszak, S.M., Habegger, L., Rozowsky, J., Shi, M., Urban, A.E., et al. (2010). Variation in Transcription Factor Binding Among Humans. Science *328*, 232–235.

Klann, T.S., Black, J.B., Chellappan, M., Safi, A., Song, L., Hilton, I.B., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2017). CRISPR-Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. Nat Biotechnol *35*, 561–568.

Liu, Y., Asakura, M., Inoue, H., Nakamura, T., Sano, M., Niu, Z., Chen, M., Schwartz, R.J., and Schneider, M.D. (2007). Sox17 is essential for the specification of cardiac mesoderm in embryonic stem cells. Proc National Acad Sci *104*, 3859–3864.

Makita, R., Uchijima, Y., Nishiyama, K., Amano, T., Chen, Q., Takeuchi, T., Mitani, A., Nagase, T., Yatomi, Y., Aburatani, H., et al. (2008). Multiple renal cysts, urinary concentration defects, and pulmonary emphysematous changes in mice lacking TAZ. Am J Physiology Ren Physiology *294*, F542-53.

Morin-Kensicki, E.M., Boone, B.N., Howell, M., Stonebraker, J.R., Teed, J., Alb, J.G., Magnuson, T.R., O'Neal, W., and Milgram, S.L. (2005). Defects in Yolk Sac Vasculogenesis, Chorioallantoic Fusion, and Embryonic Axis Elongation in Mice with Targeted Disruption of Yap65. Mol Cell Biol *26*, 77–87.

Nabet, B., Roberts, J.M., Buckley, D.L., Paulk, J., Dastjerdi, S., Yang, A., Leggett, A.L., Erb, M.A., Lawlor, M.A., Souza, A., et al. (2018). The dTAG system for immediate and target-specific protein degradation. Nat Chem Biol *14*, 431–441.

Naito, A.T., Shiojima, I., Akazawa, H., Hidaka, K., Morisaki, T., Kikuchi, A., and Komuro, I. (2006). Developmental stage-specific biphasic roles of Wnt/beta-catenin signaling in cardiomyogenesis and hematopoiesis. Proc National Acad Sci *103*, 19812–19817.

Nishizaki, S.S., and Boyle, A.P. (2017). Mining the Unknown: Assigning Function to Noncoding Single Nucleotide Polymorphisms. Trends Genet *33*, 34–45.

Noseda, M., Peterkin, T., Simões, F.C., Patient, R., and Schneider, M.D. (2011). Cardiopoietic Factors. Circ Res *108*, 129–152.

Ota, M., and Sasaki, H. (2008). Mammalian Tead proteins regulate cell proliferation and contact inhibition as transcriptional mediators of Hippo signaling. Development *135*, 4059–4069.

Paramasivam, M., Sarkeshik, A., Yates, J.R., Fernandes, M.J.G., and McCollum, D. (2011).

Angiomotin family proteins are novel activators of the LATS2 kinase tumor suppressor. Molecular Biology of the Cell *22*, 3725–3733.

Ragni, C.V., Diguet, N., Garrec, J.-F.L., Novotova, M., Resende, T.P., Pop, S., Charon, N., Guillemot, L., Kitasato, L., Badouel, C., et al. (2017). Amotl1 mediates sequestration of the Hippo effector Yap1 downstream of Fat4 to restrict heart growth. Nat Commun *8*, 14582.

Soysa, T.Y. de, Ranade, S.S., Okawa, S., Ravichandran, S., Huang, Y., Salunga, H.T., Schricker, A., Sol, A. del, Gifford, C.A., and Srivastava, D. (2019). Single-cell analysis of cardiogenesis reveals basis for organ-level developmental defects. Nature *572*, 120–124.

Totaro, A., Castellan, M., Battilana, G., Zanconato, F., Azzolin, L., Giulitti, S., Cordenonsi, M., and Piccolo, S. (2017). YAP/TAZ link cell mechanics to Notch signalling to control epidermal stem cell fate. Nat Commun *8*, 15206.

Ueno, S., Weidinger, G., Osugi, T., Kohn, A.D., Golob, J.L., Pabon, L., Reinecke, H., Moon, R.T., and Murry, C.E. (2007). Biphasic role for Wnt/beta-catenin signaling in cardiac specification in zebrafish and embryonic stem cells. Proceedings of the National Academy of Sciences *104*, 9685–9690.

Wu, S., Liu, Y., Zheng, Y., Dong, J., and Pan, D. (2008). The TEAD/TEF family protein Scalloped mediates transcriptional output of the Hippo growth-regulatory pathway. Dev Cell *14*, 388–398.

Zaidi, S., Choi, M., Wakimoto, H., Ma, L., Jiang, J., Overton, J.D., Romano-Adesman, A., Bjornson, R.D., Breitbart, R.E., Brown, K.K., et al. (2013). De novo mutations in histone-modifying genes in congenital heart disease. Nature *498*, 220–223.