MOLECULAR DISSECTION OF HAND2 DURING THE FORMATION OF PACEMAKER-LIKE MYOCYTES DURING DIRECT REPROGRAMMING

APPROVED BY SUPERVISORY COMMITTEE

Nikhil V Munshi, MD Ph.D.

Chun-Li Zhang, Ph.D.

Ondine Cleaver, Ph.D.

Eric Olson, Ph.D.

DEDICATION

I dedicate this dissertation to my future wife Christina, my mother Maricela, my dad Antonio, my three brothers, Alejandro, Abraham, and Adrian who have supported me in this chapter of my life.

MOLECULAR DISSECTION OF HAND2 DURING THE FORMATION OF PACEMAKER-LIKE MYOCYTES DURING DIRECT REPROGRAMMING

by

ANTONIO FERNANDEZ-PEREZ

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences
The University of Texas Southwestern Medical Center at Dallas
In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY
The University of Texas Southwestern Medical Center at Dallas
Dallas, Texas
May, 2019

The chapters of this dissertation have appeared in separate publications or as part of publications:

1. **Fernandez-Perez, A**, Sathe, A. A., Xing, Chao, Munshi, N.V. Hand2 selectively reorganizes chromatin accessibility to induce pacemaker-like transcriptional reprogramming. Cell Reports. 2019 Under revision.

2. Lam, H.K.*, **Fernandez-Perez, A**.*, Munshi, N. V., Schmidtke, D. W. Cyto-PDMS: A Polydimethysiloxane-based microfluidci device for intracellular cargo delivery. Biomedical Microdevices. DOI: 10.1007/S10544-018-0292-6 (2018). *Co-primary authors.

3. Sutcliffe, M. D., Tan. P. M., **Fernandez-Perez, A**., Nam, YJ., Munshi, N. V., Saucerman, J. J. High content analysis identifies unique morphological features of reprogrammed cardiomyocytes. Scientific Reports. DOI: 10.1038/s41598-018-19539-z (2018).

4. **Fernandez-Perez, A**., Munshi, N. V. Assessing Cardiomyocyte Subtypes Following Transcription Factor-mediated Reprogramming of Mouse Embryonic Fibroblasts. J. Vis. Exp. (121), e55456, DOI:10.3791/55456 (2017).

ACKNOWLEDGEMENTS

I would first like to thank my Ph.D. mentor, Dr. Nikhil Munshi, for opening the door to his laboratory and his science. This Ph.D. dissertation could not have been conceived, debated, inspired, and written without his mentorship and guidance throughout all these years. Thank you for allowing me to be creative, to explore with liberty the corners of science, and for giving me the tools to stand tall, no matter where my future takes me.

I would also like to thank the current and former members of the Munshi laboratory, for their infallible support during my time at UT Southwestern. Minoti, for guiding me when I first joined the lab. Christina, for teaching me the technique that would define my Ph.D. career. And Sam, and Lin for always having time the time to listen to my crazy ideas. Together, we shared laughs, drinks, and music that will never stop playing.

Without their wisdom and advice throughout all these years, I would had not been here today writing all my ideas and successes. Ondine, Chun-Li, and Eric, thank you for support and encouragement. Thank you—my committee.

I would like to thank my friends, Tony, Dominic, Ricky, Martyna, Varsha, Upa, Hema, Victor, Courtney, Sam, Kayla, and so many others, for all memories we created and the companionship we developed. I will always carry them with me. For this, I will be forever grateful.

Without my family I would had never had come so far from home. Their support encouraged me to leave my comfort zone and take a chance 1358 miles away from them, for an endeavor with no concrete finish line. Thank you.

And lastly, but not least, I would like to thank my future wife, Christina. I met her during my first months into my doctoral degree and she has never left my side since. Her love and support gave me the strength and motivation to push my boundaries and be better each day. This accomplishment is yours as it is mine.

Thank you all.

MOLECULAR DISSECTION OF HAND2 DURING THE FORMATION OF PACEMAKER-LIKE MYOCYTES DURING DIRECT REPROGRAMMING

Publication No._____

ANTONIO FERNANDEZ-PEREZ

The University of Texas Southwestern Medical Center at Dallas, 2019

Mentor: Nikhil, V. Munshi, M.D., Ph.D.

Direct reprogramming of one cell type into another has great promise for regenerative medicine, disease modeling, and lineage specification. Currently, the conversion of fibroblasts into induced cardiomyocytes (iCM) by Gata4, Mef2c, and Tbx5 (GMT) represents an important avenue for generating *de novo* cardiac myocytes. Recent evidence has shown that iCM formation and diversity can be enhanced by the addition of Hand2 to GMT (GHMT). These four transcription factors give rise to a heterogenous CM population, consisting of atrial (iAM), ventricular (iVM), and pacemaker myocytes (iPM). However, the molecular mechanisms that drive this plastic fate conversion remain poorly understood. Although chromatin and single-cell studies in GMT-iCM have shown the existence of a set of temporal steps that orchestrate iCM formation, little is known about how Hand2 enhances this process. In the present study, we seek to characterize these Hand2-dependent mechanisms. We hypothesize that Hand2 regulates a discrete pacemaker regulatory network that becomes active during GHMT-iCM reprogramming. To test this, we compared the transcriptional and genomic profiles of fibroblasts, GMT, GHMT, and endogenous mouse Pacemaker cells. We observe similar chromatin

landscape and gene expression profiles between Hand2-iPM and endogenous sinoatrial node (SAN), however several known key PM pathways are not active. Activation of these networks further enhances iCM-iPM fo Moreover, we show that Hand2 enhances chromatin accessibility in regions related to sarcomere function and electrical coupling, as well as promoting the closing of regions related to alternative fates. Utilizing integrative genomics between ATAC-seq and RNA-seq datasets, we identify the desmosome machinery as an important feature of iPM formation. In parallel, we define a novel Hand2 domain region that regulates cardiac subtype diversity. Taken together, our results showcase Hand2-dependent mechanisms for iPM formation and gives insight into the improvement of future iPM engineering.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

1. **Fernandez-Perez, A.,** Sathe, A. A., Xing, Chao, Munshi, N. V. 2019. Hand2 selectively reorganizes chromatin accessibility to induce pacemaker-like transcriptional reprogramming. *Cell Reports*. In Revisions**.**

2. Bhattacharyya, S., Duan, J., Wang, L., Li, B., Bhakta, M., **Fernandez-Perez, A.,** Hon, G., Munshi, V. N. 2019. Using Gjd3-CreEGFP mice to examine atrioventricular node morphology and composition. Sci Re. Accepted manuscript.

3. Lam, H.K.*, **Fernandez-Perez, A., *** Schmidtke, D.W. and Munshi, N.V., 2018. Functional cargo delivery into mouse and human fibroblasts using a versatile microfluidic device. Biomedical microdevices, 20(3), p.52. *Primary Co-authors

4. Sutcliffe, M.D., Tan, P.M., **Fernandez-Perez, A.,** Nam, Y.J., Munshi, N.V. and Saucerman, J.J., 2018. High content analysis identifies unique morphological features of reprogrammed cardiomyocytes. Scientific reports, 8(1), p.1258.

5. **Fernandez-Perez, A.** and Munshi, N.V., 2017. Assessing Cardiomyocyte Subtypes Following Transcription Factor-mediated Reprogramming of Mouse Embryonic Fibroblasts. Journal of visualized experiments: JoVE, (121).

6. Nam, Y.J., Lubczyk, C., Bhakta, M., Zang, T., **Fernandez-Perez, A.**, McAnally, J., Bassel-Duby, R., Olson, E.N. and Munshi, N.V., 2014. Induction of diverse cardiac cell types by reprogramming fibroblasts with cardiac transcription factors. Development, pp.dev-114025.

7. **Fernandez, A**., Huggins, I.J., Perna, L., Brafman, D., Lu, D., Yao, S., Gaasterland, T., Carson, D.A. and Willert, K., 2014. The WNT receptor FZD7 is required for maintenance of the pluripotent state in human embryonic stem cells. Proceedings of the National Academy of Sciences, 111(4), pp.1409-1414.

8. Mahadevan, N.R., Rodvold, J., Almanza, G., **Pérez, A.F**., Wheeler, M.C. and Zanetti, M., 2011. ER stress drives Lipocalin 2 upregulation in prostate cancer cells in an NF-κB-dependent manner. BMC cancer, 11(1), p.229.

9. Brafman, D.A., Chang, C.W., **Fernandez, A**., Willert, K., Varghese, S. and Chien, S., 2010. Long-term human pluripotent stem cell self-renewal on synthetic polymer surfaces. Biomaterials, 31(34), pp.9135-9144.

10. Almanza, G., **Fernandez, A.,*** Volinia, S., Cortez-Gonzalez, X., Croce, C.M. and Zanetti, M., 2010. Selected microRNAs define cell fate determination of murine central memory CD8 T cells. PloS one, 5(6), p.e11243. *Primary Co-authors

11. Mahadevan, N.R., **Fernandez, A.,** Rodvold, J.J., Almanza, G. and Zanetti, M., 2010. Prostate cancer cells undergoing ER stress in vitro and in vivo activate transcription of pro-inflammatory cytokines. Journal of inflammation research, 3, p.99.

LIST OF FIGURES AND TABLES

CHAPTER ONE

CHAPTER TWO

CHAPTER THREE

CHAPTER FOUR

CHAPTER FIVE

LIST OF APENDICES

LIST OF ABBREVIATIONS

AVN- Atrioventricular node

bHLH- Basic helix-loop-helix

CCS- Cardiac Conduction System

CHD- Congenital Heart Disease

CM- Cardiomyocyte

cTnT- Troponin T2

FACS: Fluorescent activated cell sorting

FHF- First Heart Field

Gata4- GATA binding protein 4

GHMT- Gata4, Mef2c, Tbx5, and Hand2

GMT- Gata4, Mef2c, and Tbx5

Hand1- Heart and neural crest derivatives expressed transcript 1

Hand2- Heart and neural crest derivatives expressed transcript 2

iALM- Induced atrial-like myocyte

ICC- Immunocytochemistry

iCLM- Induced cardiomyocyte-like myocyte

iCM- Induced cardiac myocyte cell

IHD- Ischemic heart disease

iPLM- Induced pacemaker-like myocyte

iPM- Pacemaker myocyte

iVLM- Induced ventricular-like myocyte

iVM-Ventricular myocyte

LA- Left Atria

LV- Left Ventricle

MEF(s)- Mouse embryonic fibroblasts

Mef2c- Myocyte enhancer factor 2c

NGS- Next Generation Sequencing

OTF- Outflow track

Plat-E- Platinum-E

RA- Right Atria

RV- Right Ventricle

SAN- Sinoatrial node

SHF- Second Heart Field

SHH- Sonic Hedgehog

Tbx5- T-box 5

TF- Transcription factor

α-actinin- Alpha-actinin protein

CHAPTER ONE
GENERAL INTRODUCTION

Chapter overview

Every year, approximately seven million individuals worldwide die of Ischemic heart disease (IHD), which has emerged as the leading cause of death(Go et al., 2014). Therefore, understanding cardiac regeneration and repair has become a vital aspect of current cardiovascular research. Unfortunately, the adult mammalian heart has a limited regenerative potential(Lin and Pu, 2014). Thus, the need to find alternative therapies to replace injured adult myocardium is of critical importance.

Recent studies have demonstrated the potential of generating *de novo* cardiomyocytes (CM) from dermal fibroblasts. Their potential as a source of CM, provides opportunities to repair the heart. However, many challenges and roadblocks lie ahead. This thesis focuses on some of those challenges and will try dissecting how one transcription factor is able to modulate cell fate. But first, in this chapter, I would introduce basics concepts that I believe will aid us in understanding the complexities, challenges, and the impact of my work. We will start with an overview of the major stages of cardiac development, followed by details of the conduction system, and the nuances of the structural complexities of a cardiomyocyte. In the last sections, we will talk about bHLH transcription factors, their role in heart, and lastly, I will give an overview on the different direct reprogramming strategies currently available.

Cardiac development

The embryonic heart is the first organ to function during development. Any insult that interferes with the complex gene regulatory networks that regulate its development can lead to congenital heart diseases (CHD), of the major causes of mortality in infants and adults(Bruneau, 2008). In the mouse embryo, the heart can be first observed as a crescent-shape structure at embryonic day E7.5(Xin et al., 2013). This cardiac crescent is composed of myocardial progenitors from the first and second heart fields that will contribute to different aspect of the adult heart. The myocardial progenitors form a simple tube that begins beating at around E8.0. This heart tube is composed mainly by first heart field (FHF) cells and it will serve as a stage for the secondary heart fields (SHF) cells to migrate from the adjacent splanchnic pharyngeal mesoderm to the venous and arterial poles of the tube (**Figure 1.1).** The venous pole arises from myocardial cells stemming from the dorsal myocardium, early superior vena cava and pulmonary vein; is the entry point of the blood in the heart tube. The arterial pole forms at the base of the aorta, is where the blood exits the heart tube. The outflow track (OFT), a transient structure connects the right ventricle to aortic sac. At E8.5, the heart begins to loop. It first

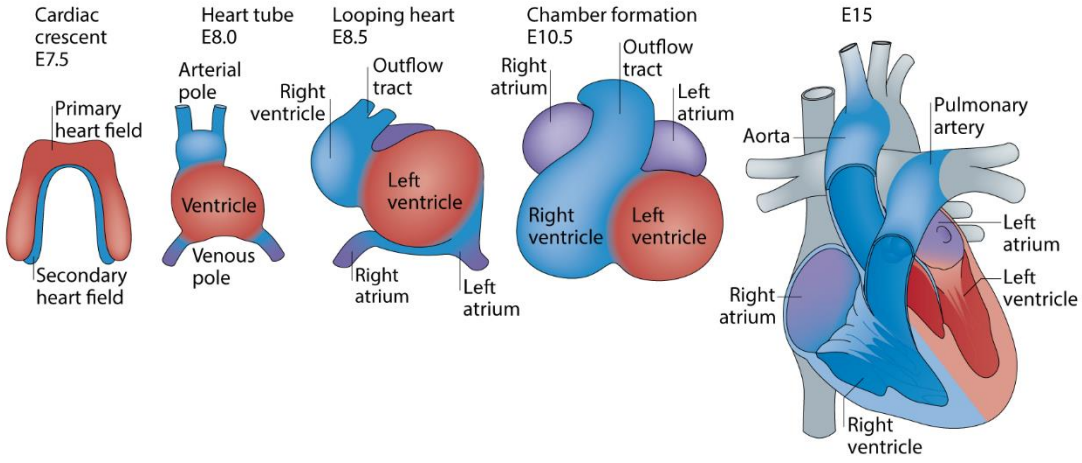


Figure 1.1 Cardiac development in the mouse. *From left to right: cardiac crescent stage (E7.5), heart tube stage (E8.0), loping heart (E8.5), formation of the chambers*

(E10.5) and functional four chambers heart (E15.0). Red indicates the primary heart field (PHF) and blue the second heart the field (SHF) and their derivative structures. Adapted from (Modified from Xin et al., 2013).

undergoes a rightward looping creating the primitive ventricles and atria. From E10.5-E12, a multilayered compact zone that contributes to the thickness of the ventricular wall arises from the proliferation of the cardiomyocytes in the myocardium of the ventricles. Some of these myocytes continue to migrate deeper into the ventricles and give rise to trabeculations, which will serve as a temporal the oxygenation machinery until the coronary circulation is established in the myocardium (Martin-Puig et al., 2008). During this period, you also have the formation of the cardiac conduction system and valves(van Eif et al., 2018a). The cardiac valves develop both the OFT and the atrioventricular canal, connecting the left ventricle to the common atria. In parallel, the formation of the septa (interventricular, interatrial, and atrioventricular) divide the heart into the functional four chambers. (Lin et al., 2012).

Cardiac conduction development and Sinoatrial node

The cardiac conduction system can be divided into the slow- and fast-conducting nodes. The slow-conducting include the sinoatrial node (SAN) and the atrioventricular node (AVN). The fast-conducting includes the ventricular conduction system (VCS), the atrioventricular bundle, bundle branches (BB), and the Purkinje fiber network.

The SAN is the principal pacemaker of the heart and it determines the rate of contraction. Physically, it is located in between the right atrium (RA) and the superior vena cava. The pacing of the SAN is modulated by highly-specialized cardiomyocytes called pacemakers myocytes, which can instantaneously depolarize their membranes through

the help of potassium channels. Once the signal is initiated in the SAN, the action potential travels through the atria to the AVN, which delays the signal until the ventricles fill with blood. The electrical signal then rapidly propagates through the BB and the VCS.

The first cardiac pacemaker signals can be observed in the venous pole after cardiac tube formation at E8.5(Tyser et al., 2016). The impulse is transmitted slowly through the myocardium, followed by nascent contractions(Anderson et al., 2006). Following the formation of the primitive cardiac tube, the mesenchymal cardiac progenitors come together to form the sinus horns, sinus venous (SV), and then SAN (van Eif et al., 2018a). (**Figure 1.2**)

While the SV and SAN are differentiating, the cardiac progenitors cells that are located in the venous pole, and will give rise to sinus horn, start to express Tbx3, a member of the T-box family that marks the central conduction system(Hoogaars et al., 2007a), as well as Hcn4, a hyperpolarization-activated nucleotide-gated cation channel, highly enriched in SAN during development(Liang et al., 2013a). In murine cardiac progenitor cells, expression of Nkx2-5 is present until ~E9.0, which marks establishment of the atrial chamber(Stanley et al., 2004). From E9 to E12, the progenitors express Tbx18 instead of Nkx2-5, establishing a myocardium with Nkx2-5l^{ow}Tbx18^{+}(Christoffels et al., 2006). In parallel, Hcn4 expression is reduced in cardiac progenitors that will give rise to the atrial chambers, and will be restricted to Nkx2-5l^{ow}Tbx18^{+} sinus venosus/SAN(Wiese et al., 2009). **(Figure 1.2b-c**). The final pacemaker activity shift to the sinus venous and to the SAN-to be anatomy. The unrestricted portion of the sinus venosus begins a process

called atrialization of the SV, and myocardial genes begin to express (Gja1, Gja5, and Nkx2-5).(Mommersteeg et al., 2007b).

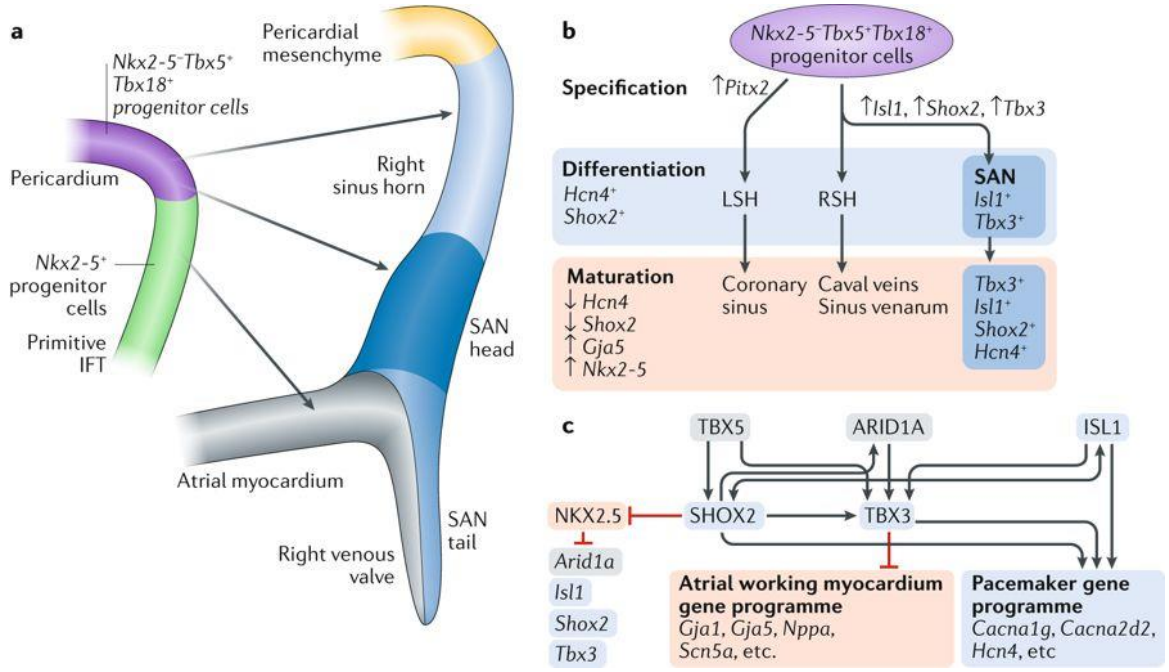


***Figure 1.2 Molecular mechanism that regulate SAN and sinus horn development**. **A)** The sinoatrial node (SAN) and sinus horn develop from $Tbx18^{+}Nk2x.5^{-}$ progenitor cells. **B) Left side:** Pitx2 expression negatively regulates development of the SAN. **Right side:** SAN progenitors express Isl1, Shox2, and Tbx3. Pacemaker activity becomes restricted to the SAN, and Hcn4 and Shox2 expression is established. **C)** Transcriptional networks in SAN development. Adapted from (van Eif et al., 2018a)).*

SAN development to the Right sinus horn is controlled by Pitx2, Pituitary homeobox 2 protein, in where suppresses the right-side SV gene expression. Isl1, Insulin gene enhancer protein, is also a critical component of the PM network. It is originally expressed by cardiac progenitor cells, and later on restricted to the SAN, where it is maintained through adulthood(van Eif et al., 2018a). It regulates Bmp4, Hcn4, Shox2, and Tbx3 expression. (Liang et al., 2015b). Tbx3 expression is tightly controlled in the SAN, at both early, and late developmental stages. In the SAN, it represses Gja1, Gja5, and Scn5a(Hoogaars et al., 2007a).

Cardiomyocyte structure and function

The sarcomere is the principal component of the cardiac myocyte and it is responsible for the contraction and expansion of the muscle cells. These structures are composed of a series of myofibrils parallel to the long axis of the myocyte. Moreover, these fibrils are composed of thick and thin filaments which connect at the Z disc. The thick filaments are composed of myosin heavy chain and light chains centered around the M line and the thin filaments are composed of actin, troponin, and tropomyosin, to name a few(Koubassova and Tsaturyan, 2011). When the contraction signal is received, calcium is released from the sarcoplasmic reticulum which binds troponin, triggering the movement of the tropomyosin and the binding of actin; this actin binding activates the myosin ATPase activity causing its movement along actin and pulling Z discs nearer, resulting in the contraction of the cell(Koubassova and Tsaturyan, 2011) (**Figure 1.3).**

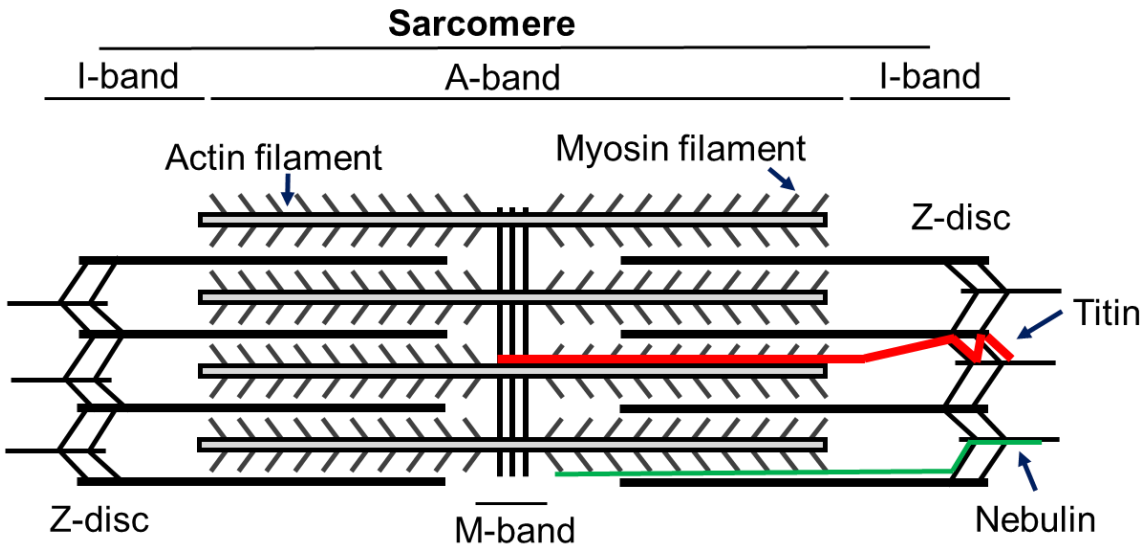


Figure 1.3 Striated muscle sarcomere. A) *Schematic diagram showing the main components of the sarcomere. The A-band comprises myosin filaments crosslinked at the center by the M-band assembly. Thin actin-containing filaments are tethered at their barbed end at the Z-disc and interdigitate with the thick filaments in the A-band. Nebulin (800 kDa) runs along the thin filaments and overlaps in the Z-disc (Pappas et al., 2008)). The 3 MDa 1 μm long protein titin runs between the M-line and the Z-disc (Young et al., 1998).*

Actin is a highly conserved globular protein found in every eukaryotic cell. It has a key role in myocyte contraction, cytoskeletal structure, morphology, protein localization. It can be found as a monomer, G-actin, or as polymer, F-actin. The 43 kDa G-actin monomer is composed of 375 residues and it polymerizes to form filamentous F-actin(Kabsch et al., 1990). These F-actin filaments elongate at the (+) end and shorten at the (-) end unless they are regulated by CapZ, which prevents the addition of actin at the (+) end(Vavylonis et al., 2005). Like actin, α-actinin is ubiquitous to eukaryotic cells, it belongs to the spectrin family of actin-binding proteins which include spectrin, dystrophin, utrophin and fimbrin. In muscle cells, it intercalates with Z-discs between actin filaments of alternative polarity from adjoining sarcomeres. It is comprised of an N-terminal acting binding domain (ABD), four spectrin-like repeats (R1-R4) comprising triple helical coiled-coils that form a ~35 nm rod shape domain and two EF hand domains (**Figure 1.4).** Using cryo-electron microscopy McGough et al. (McGough et al., 1994) showed that a-actinin binds over two neighboring actin monomers along the long helix (through residues 348-355 and 87-96). In addition to its structural role other studies have implicated actinin with docking of signaling proteins at the Z-disc(Sjöblom et al., 2008).

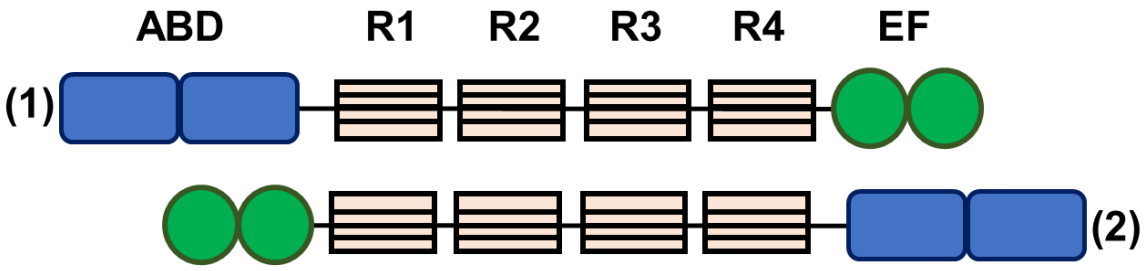


Figure 1.4 Cartoon structure of α-actinin. A) *α-actinin contains and N-terminal actin domain (ABD), four spectrin-like repeats (R1-R4) comprising triple helical coiled-coils that form a rod shape domain, and two EF hand domains (calmodulin homology domains). Cartoon depicting an α-actinin dimer (1) and (2).*

Moreover, we have previously shown that proper sarcomere organization is a prerequisite of functional reprogrammed cardiomyocytes(Nam et al., 2014), the dynamics of how each of the components mentioned above should then be topic of interest of direct-CM reprogramming field.

Overview of bHLH transcription factors

The basic helix-loop-helix (bHLH) are part of a large superfamily of transcriptional regulators that play a crucial role in the specification and differentiation of a multitude of tissues and organs during development. Using methylation protection assays, the first HLH domains identified were associated with IgH and B cell enhancer studies(Ephrussi et al., 1985). Since these bHLH factors bind to Ephrussi-box (E-box) sequences (CANNTG), they were first named E proteins(Murre et al., 1989). Interestingly, only one E protein has been identified in the *Drosophila,* the daughterless (Da) protein.

Members of the bHLH superfamily have two highly conserved domains, making up a region of approximately 60 amino-acid residues. At the N-terminus end of this regions there is basic domain which bind the transcription factor to E-box sequences. Different bHLH family members are known recognize different E-box consensus sequences. Murre et al(Jones, 2004; Murre et al., 1994), proposed their classification based on sequence comparisons and is summarized in **Table 1.1** (Adapted from Jones, 2004). The diversity of the E-box sequences and the dimers formed by the bHLHs suggests tight regulated control of transcriptional regulation.

The first bHLH motif was identified by Murre et al., in two murine TF, E12 and E47. These E proteins have a wide variety of functions as heterodimer partners of the tissue-specific bHLHs proteins, including neurogenesis, cardiogenesis, and myogenesis. For example, NeuroD, and Ascl1 (Lee, 1997) has been shown to be involved in terminal neuron differentiation. Even more recently, ASCL1 has been used to reprogrammed fibroblasts into induced Neuronal cells(Chanda et al., 2014). Similarly, the myogenic factors, MyoD, and Myf-5, and Myogenin are important for the establishment and differentiation of the myogenic lineage(Weintraub et al., 1991). MyoD alone is also able to reprogram fibroblast into skeletal muscle cells(Davis et al., 1987). The transcription factors Hand1 and Hand2 are important in cardiac development in vertebrates(Srivastava and Olson, 1997).

Table 1.1

Classification of bHLH proteins by sequence			
Phylogenic group	*Description*	*Classification according to Murre et al*	*Examples of family members*
A	Bind to CAGCTG or CACCTG	I, II	MyoD, Twist
B	Bind to CACGTG or CATGTTG	III, IV	Mad, Myc
C	Bind to ACGTG or GCGTG. Contain a PAS domain		Arnt
D	Lack a basic domain and hence do not bind DNA but V ID form protein-protein dimers that function as antagonists of group A proteins	V	ID
E	Bind preferentially to N-box sequences CACGCG or VI Hairy CACGAG. Contain an orange domain and a WRPW peptide	VI	Hairy

The Hand2 in the heart

In vitro, Hand2 interacts at the highest affinity with the CATCTG E-box ((Dai and Cserjesi, 2002). Although, data from this thesis suggest a more complex set of rules dictating the DNA binding consensus based on context (See Chapter 5). Hand2 and Hand1, as other bHLH TFs require dimerization of two bHLH proteins via their HLH domain for the E-box recognition(Vincentz et al., 2011).

In the context of the heart, Hand2 (previously dHand2) is expressed in the deciduum, heart, autonomic nervous system, and neural crest derivatives and is critical for the development of these structures. Moreover, Hand2 regulates craniofacial structures derived from neural crest cells(Barron et al., 2011) and several enzymes, like norepinephrine, that regulate cardiac rhythm(VanDusen et al., 2014). Furthermore, is essential in the formation of the posterior forelimb and hindlimb bud through SHH activation(Galli et al., 2010).

During mouse cardiac development, Hand2 expression is first detected in the SHF progenitors at E7.5. It is further maintained in the heart tube before its restriction to the right ventricle (RV) and OFT myocardium during looping; with some minimal expression in the left ventricle (LV) at E8.5. Following tube formation, Hand2 expression is detected in the endocardium and the splanchnic pharyngeal mesoderm. At 9.5 and later, Hand2 is expressed in the proepicardial organ and derived epicardium, and branchial arches. During cardiac chamber remodeling, Hand2 is restricted in the atria(Barnes et al., 2011; Dirkx et al., 2013; Srivastava et al., 1995; Srivastava and Olson, 1997; Srivastava et al., 1997). In contrast Hand1 is expressed mainly in the FHF structures and it is restricted to the LV during looping (**Figure 1.5**). Hand2 and Hand1 are known to form heterodimers,

and they redundantly in the differentiation of the FHF-LV(Firulli et al., 2010b). Firulli et al shows that Hand1 does not have a role on the formation of SHF-RV(Firulli et al., 2000).

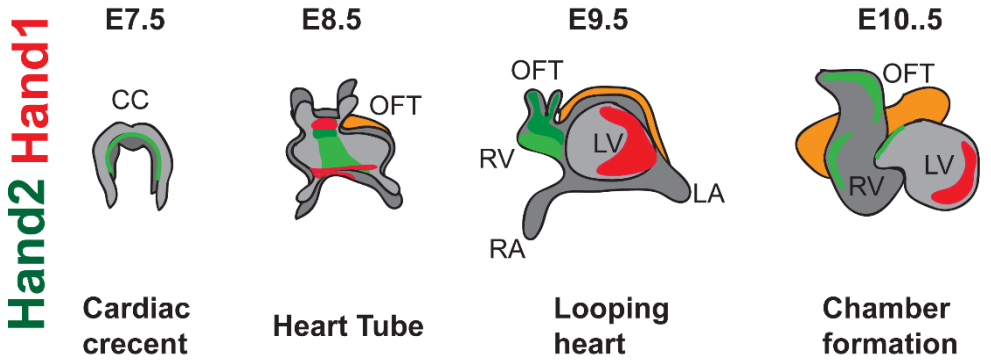


Figure 1.5 *Schematic representation of the expression of Hand1 (Top, Red) and Hand2 (Green, bottom) during development. CC: cardiac crescent. OFT: Outflow track. RV: Right Ventricle. RA: Right Atria. LA: Left Atria. LV: Left Ventricle.*

The importance of Hand2 during heart development can be best appreciated by work done by Srivastava et al. (Srivastava and Olson, 1997). By knocking out Hand2, embryos displayed severe cardiac defects associated with growth retardation and embryonic lethality by E10.5. No noticeable morphological changes were appreciated until the cardiac looping stage. After that, embryos showed signs of hypoplastic RV and branchial arches, lack of trabeculation in the myocardium, dilation of the aortic sac, among other gross changes(Srivastava and Olson, 1997; Togi et al., 2006).

To further dissect Hand2 function in the heart, Liu et al. (Liu et al., 2009) abolished the DNA-binding activity of Hand2 and showed that Hand2^{EDE} embryos show normal development of the heart and branchial arches until E11.5. Instead embryonic lethally occurred during the reduced RV growth. This suggest that Hand2 carries DNA-dependent and independent functions in heart. Although this thesis does not explore in detail this observation, chapter 5 alludes at potential explanations to these results.

To summarize, the bHLH superfamily encompasses a diverse range of proteins, with about 125 been identified in humans. Their function includes, cell cycle regulation, tissue-specificity, lineage commitment, and tumorigenesis. It is important to highlight that so far only nine bHLH protein structures have been annotated, and eight of these have been classified to a single superfamily(Jones, 2004). Still, little is known about the mechanisms that they utilize to control their wide range of processes. Nonetheless this thesis brings light, in the context of Hand2, the plastic mechanisms by which bHLH can be functioning to regulate gene expression.

Direct reprogramming of cardiac myocytes

The seminal study by Ieda et al(Ieda et al., 2010a) showed that dermal fibroblast (as well as postnatal cardiac fibroblasts) were capable of being directly reprogrammed into cardiomyocyte-like cells in vitro by the addition of three cardiac-associated transcription factors Gata4, Mef2c, and Tbx5 (GMT). In subsequent studies, Qian et al(Qian et al., 2012) used GMT to in vivo reprogramed murine cardiac fibroblast and recover cardiac function after injury. In 2012, Song et al(Song et al., 2012), showed you can improve the efficiency of in vitro reprogramming by the addition of the transcription factor Hand2 to GMT (GHMT). Nam et al(Nam et al., 2014) subsequently showed that GHMT is able to generate a wide variety of cardiac subtypes, bringing to light the plasticity of the system and the challenges of targeting a single cardiac subtype.

Following these original reports, different groups approach this paradigm from different angles. With clinical application in mind, Lee and Li (Lee et al., 2015; Li et al., 2015) showed that is possible to deliver mRNAs and purified cardiac transcription factors

to induce iCM and cardiac progenitor cells with varying degrees of efficiency. With the inclusion of gene expression analysis and later next-generation sequencing (NGS) different reprogramming roadblocks were identified and several groups develop novel cocktails to improve efficiencies. Both Ifkovits and Zhao (Ifkovits et al., 2014b; Zhao et al., 2015) showed that by inhibiting pro-fibrotic signaling (SB431542 for TTGF-β and iRock for Rho-associated kinase pathways) reprogramming efficiencies could reach 60% for expression for cTnT. Furthermore, Yamakawa et al(Yamakawa et al., 2015b) modulated p38MAPK and PI3K/AKT pathways to increase activation of cardiac transcription regulators increasing reprograming efficiencies and most importantly beating frequency. With a shRNA screen, Zhou et al identified Bmi1 as a critical epigenetic barrier for reprograming(Zhou et al., 2016). Similarly, Zhou et al(Zhou et al., 2015) improved cardiac efficiencies and maturation with the addition of Akt1 and later identifying inflammation as a roadblock for reprograming and repressing it with the zinc finger, ZNF281(Zhou et al., 2017). Moreover, Qian Li and colleagues(Liu et al., 2017b; Wang et al., 2015) show the importance of transcription factor stoichiometry and managed to deconstruct GMT-iCM transcriptome at a single cell resolution, brining into light other previously unknown reprogramming roadblocks.

While the field has made great advances in improving reprogramming efficiencies (Summary of current approaches **Table 1.2**), there are still many challenges ahead. One of the most important one is bringing standardization to the output assay. The cardiomyocyte is a complex cell, composed of a multitude of structures of different nature that comes together to perform a function (**Figure 1.6**), and it is important that these major components are being taken into consideration before determining cardiac efficiencies.

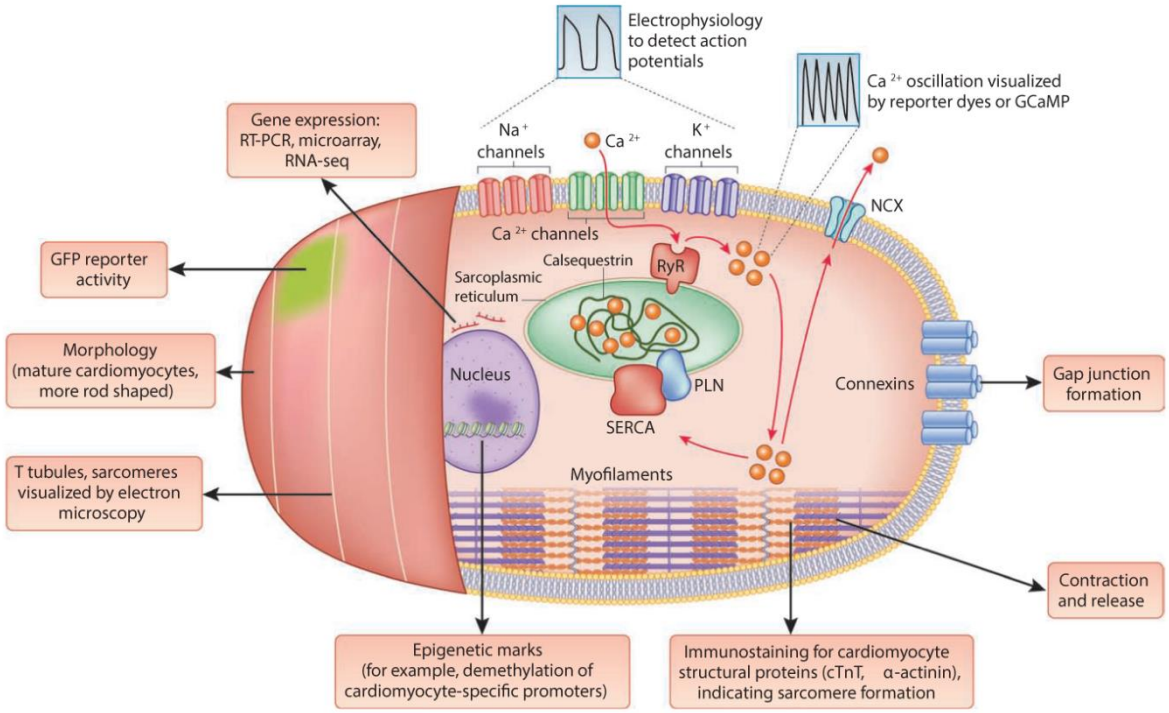


Figure 1.6 How to make a cardiomyocyte. *A cardiomyocyte is a complex cell with many functional components. These include: activation of a reporter transgenes (such as Hcn4–GFP) to more complex functional characteristics, such as the ability to fire action potentials and show calcium oscillation and action potentials. Other testable characteristics include the presence of cardiomyocyte-specific epigenetic marks, structural characteristics, gap junctions, desmosome machinery, nuclear morphology, among others. T tubule, transverse tubule; RyR, ryanodine receptor; SERCA, sarco/endoplasmic reticulum Ca2+-ATPase; NCX, sodium-calcium exchanger; PLN, phospholamban; GCaMP, a genetically encoded calcium indicator consisting of a fusion of GFP and calmodulin (CaM). Modified from(Addis and Epstein, 2013).*

Lastly, there should be a comprehensive effort in documenting the differences that media components, viral titers, cardiac subtypes outputs, starting cell population properties, delivery constructs properties (retro-/lentivirus/AAV/retrotransposons), etc. In short, direct reprogramming is a sensitive and complex process that will require more study.

Table 1.2. Modified from(Chen et al., 2017)

Combination of factors	Source	Markers and efficiency	AP	Beating	Group
GMT	CF, TTDF	cTnT$^+$: 30% of α-MHC cells; α-actinin$^+$: most of cTnT$^+$ cells	+	+	(Ieda et al., 2010a)
OSKM; JI1, BMP4	MEF	cTnT$^+$: 40%	+	+	(Efe et al., 2011)
GMT	CF	α-MHC-EYFP$^+$: ~40% at border zone	+	+	(Qian et al., 2012a)
miR-1,*133*,*208*,*499*; JAK inhibitor I	CF	α-MHC-GFP$^+$: ~28%	+	+	(Jayawardena et al., 2012)
GMT, *Myocd*, *Srf*, *Mesp1*, *Smarcd3*	MEF	Myh6.Egfp$^+$: 2.4%	–	–	(Christoforou et al., 2013)
Hand2, *Nkx2.5*, *Gata4*, *Mef2c*, *Tbx5*	MEF, CF	Troponin T-GCaMP5$^+$ activity: 1.6%	ND	+	(Addis et al., 2013)
GHMT, *MyoD* domain	HF, LBF, TTF	cTnT$^+$: 4.9%	ND	+	(Hirai et al., 2013)
GHMT and SB431542	CF	Troponin T-GCaMP5$^+$ activity: 9.27%	ND	–	(Ifkovits et al., 2014b)
GHMT, *Myod* domain	HF	cTnT$^+$: 19%	ND	+	(Hirai and Kikyo, 2014)
GMT, *Mesp1*, *Myocd* and miR-133	MEF, CF	α-MHC-GFP & cTnT$^+$: 8.1%; α-actinin$^+$: 19.9%	ND	–	(Muraoka et al., 2014b)
OCT4, SCPF	MEF, TTF	beating clusters:~40/well of 24-well plate	+	+	(Wang et al., 2014a)
GHMT	MEF,	Sarcomere$^+$: ~32%; NPPA$^+$: 35% of sarcomere$^+$; MYL2$^+$: 22% of sarcomere$^+$	+	+	(Nam et al., 2014b)
GMT mRNA, C_lipo	CF	α-MHC-GFP$^+$: 0.5% of transfected CF	ND	–	(Lee et al., 2015)
miR-1, *miR-133*, *miR-208*, *miR-499*	CF	tdTomato$^+$ Troponin T$^+$:12%	+	+	(Jayawardena et al., 2012)
OSKM, Ascorbic acid	MEF	GATA4$^+$: ~40%; MHC$^+$: ~24%	+	+	(Talkhabi et al., 2015)
CHIR99021, RepSox, Forskolin, VPA	MEF, TTF	α-actinin$^+$: 14.5%; α-MHC$^+$: 9%	+	+	(Fu et al., 2015)
GHMT, *miR-1*, *miR-133*, Y-27632, A83-01	MEF, AF	cTnT$^+$: ~60% with A83-01; α-actinin$^+$: ~60% with A83-01	+	+	(Zhao et al., 2015)
GHMT + AKT	MEF, CF, TTF	~27% cTnT$^+$α-MHC in MEFs	–	+	(Zhou et al., 2015)
AGHMT + ZFN281	TTF	~27% cTnT$^+$α-MHC in TTFs	–	–	(Zhou et al., 2017)

CHAPTER TWO

A METHODLOGY FOR CARDIAC SUBTYPE QUANTIFICATION

The contents of this chapter are reproduced from:

> **Fernandez-Perez, A**. and Munshi, N.V., 2017. Assessing Cardiomyocyte Subtypes Following Transcription Factor-mediated Reprogramming of Mouse Embryonic Fibroblasts. Journal of visualized experiments: JoVE, (121).

Chapter Overview

In the previous chapter, I reviewed basic principles of heart development, the diversity of non-cardiac cells in the heart, and an overview of the status of direct cardiac reprogramming. In this second chapter, I will focus on the methodology that I used to generate all the data during my doctoral thesis career. Given the intricacies of direct reprogramming, I present a detail visual protocol for the reprogramming of mouse embryonic fibroblasts into the diverse cardiac subtypes. Due to the importance of cell type identification, I will go over in detail the methodology that I used to manually quantify iCM-Subtypes. As a personal anecdote, one experiment of n= 24 samples (12mm coverslip) would take approximately 16 hr of confocal microscopy to quantify and tabulate.

For this publication, I wrote the optimized iCM protocol, carried out the reprogramming experiments, the quantification, edited the video script, and performed the visuals. Please see related video on JoVE website for better representation of the technical procedures.

ABSTRACT

Direct reprogramming of one cell type into another has recently emerged as a powerful paradigm for regenerative medicine, disease modeling, and lineage specification. In particular, the conversion of fibroblasts into induced cardiomyocyte-like myocytes (iCLMs) by Gata4, Hand2, Mef2c, and Tbx5 (GHMT) represents an important avenue for generating *de novo* cardiac myocytes *in vitro* and *in vivo*. Recent evidence suggests that GHMT generates a greater diversity of cardiac subtypes than previously appreciated, thus underscoring the need for a systematic approach to conducting additional studies. Before direct reprogramming can be used as a therapeutic strategy, however, the mechanistic underpinnings of lineage conversion must be understood in detail to generate specific cardiac subtypes. Here we present a detailed protocol for generating iCLMs by GHMT-mediated reprogramming of mouse embryonic fibroblasts (MEFs).

We outline methods for MEF isolation, retroviral production, and MEF infection to accomplish efficient reprogramming. To determine the subtype identity of reprogrammed cells, we detail a step-by-step approach for performing immunocytochemistry on iCLMs using a defined set of compatible antibodies. Methods for confocal microscopy, identification, and quantification of iCLMs and individual atrial (iAM), ventricular (iVM), and pacemaker (iPM) subtypes are also presented. Finally, we discuss representative results of prototypical direct reprogramming experiments and highlight important technical aspects of our protocol to ensure efficient lineage conversion. Taken together, our optimized protocol should provide a stepwise approach for investigators to conduct

meaningful cardiac reprogramming experiments that require identification of individual CM subtypes.

INTRODUCTION

The heart is the first functional organ to develop in the embryo(Buckingham et al., 2005; Sissman, 1970). In conjunction with the circulatory system, it supplies oxygen, nutrients, and a waste disposal mechanism during development. Three weeks after fertilization, the human heart beats for the first time and its proper regulation is maintained by cardiomyocytes (CMs). The irreversible loss of these specialized cells is therefore the fundamental issue underlying progressive heart failure. While some organisms such as the zebrafish and Xenopus have the potential for cardiac regeneration, the adult mammalian heart is more limited(Ali et al., 2014; Bergmann et al., 2009; Lin and Pu, 2014; Senyo et al., 2013). Thus, given the critical function of the heart, it is not astonishing that heart disease is the leading cause of death in the world, accounting for 600,000 deaths in the United States alone(Writing Group et al., 2016). Therefore, cell-based therapies to efficiently repair or replace the injured myocardium are of great clinical interest.

The seminal study lead by Yamanaka and colleagues(Takahashi and Yamanaka, 2006) showed that forced expression of four transcription factors is sufficient to convert fully differentiated fibroblast cells to pluripotent stem cells. However, the tumorigenic capacity of all pluripotent stem cell strategies has been a critical concern in their use for therapeutic purposes. This motivated the scientific field to search for alternative methods to transdifferentiate cells while avoiding a pluripotent stage. Recently, several groups

have shown the feasibility of this strategy by displaying direct conversion of mouse fibroblasts to induced cardiomyocyte-like cells (iCLMs) with the ectopic expression of the transcription factors Gata4, Mef2c, Tbx5, and later on, Hand2 (GMT and GHMT, respectively)(Ieda et al., 2010b; Song et al., 2012b). Subsequent studies showed that the same strategy could be performed *in vivo* and in human-derived tissues *(Fu et al., 2013b; Qian et al., 2012b; Song et al., 2012b)*. Even more recently, new studies have shown that this process is furthered improved by targeting different aspects of the reprogramming paradigm(Zhao et al., 2015; Zhou et al., 2015; Zhou et al., 2016).Taken together, these studies demonstrate the potential of directed transdifferentiation for regenerative therapies. However, the low efficiency of CM reprogramming, the unknown molecular mechanisms, inconsistent reproducibility due to methodological differences(Miki et al., 2013), and the heterogeneous nature of iCMs remain unaddressed.

To directly evaluate iCM heterogeneity, we designed a discrete and robust single-cell assay for the identification of sarcomere development and cardiac lineage specification–two necessary characteristics of functional cardiomyocytes. There are at least three major types of CM in the heart as defined by their location and unique electrical properties: atrial (AM), ventricular (VM) and pacemaker (PM)(Atkinson et al., 2011; Bootman et al., 2011; Miquerol et al., 2011; Spater et al., 2014). In an orchestrated combination, they allow the proper pumping of the blood. During heart injury, one or all subtypes might be affected, and the type of cell therapy would need to be addressed on a case-by-case basis. Currently, most strategies are focusing on the overall generation of cardiomyocytes, while little work is being done in studying the molecular mechanism that regulates subtype specification.

The following study details how to properly quantify well-organized sarcomeres and identify a diverse set of cardiomyocyte subtypes. Using a pacemaker (PM)-specific reporter mouse, we can apply an immunocytochemical approach to distinguish induced atrial-like myocytes (iAM), induced ventricular-like myocytes (iVM), and induced PM-like myocytes (iPMs)(Nam et al., 2014b). Based on our observations, only cells that exhibit a particular degree of sarcomere organization are the ones capable of spontaneous beating. This unique reprogramming platform allows for assessing the role of certain parameters in sarcomere organization, subtype specification, and efficiency of CM reprogramming at single-cell resolution.

PROTOCOL

All experimental procedures involving animal practices were approved by the Institutional Animal Care and Use Committee at UT Southwestern Medical Center.

1. Isolation of Hcn4-GFP E12.5 mouse embryonic fibroblast (MEFs)

1.1) Set up timed matings between homozygous Hcn4-GFP males and CD-1 females.

1.2) Sacrifice pregnant female at E12.5 by carbon dioxide euthanasia, followed by cervical dislocation.

1.2.1) Remove uterine horns with dissecting forceps and place them into a petri dish on ice with 1x PBS without $Ca^{2+}Mg^{2+}$ (Conner, 2001; Jozefczuk et al., 2012).

1.3) Perform all subsequent steps in the tissue culture hood using sterile technique.

1.3.1) Remove the embryos from the uterus and amniotic sac using scissors and dissecting forceps. Keep the placenta attached for better handling. Homozygous

Hcn4-GFP pregnant females give birth between 10-14 pups.

1.3.2) Using dissecting forceps take the isolated embryos and briefly rinse them twice in 70% (v/v) EtOH. Note: the washes should be fast to minimize cell death.

1.3.3) Remove the head, limbs, tail and internal organs, including the heart from the isolated embryos.

1.3.4) Finely mince the remaining tissue using a sterile razor blade to about 1mm of size in a 10-cm dish with 1 mL of 1x PBS.

1.3.5) Transfer minced tissue into a 50 mL conical tube with PBS.

1.3.6) Spin at 300 x g for 3 min. Carefully aspirate excess PBS.

1.3.7) Add 1 mL of sterile 0.25% trypsin-EDTA per embryo. Incubate cells in 37 °C water bath for 15 min. Gently mix the tube every 4 min. Over digestion of the tissue leads to low yield.

1.3.8) Vortex cell mixture at maximum speed (3200rpm) for 4 s.

1.3.9) Add 2 mL of fibroblast media per embryo and mix. Filter through a 100 μm cell strainer. Use a pipette to aid the cells through the strainer. Refer to **Table 2.1** for the formulation of all subsequent mediums.

1.3.10) Spin at 300 x g for 4 min. Carefully aspirate supernatant.

1.3.11) Add 10 mL of fresh fibroblast media per every 3 embryos and triturate 6-10 times.

1.3.12) Plate the cells in a 15-cm tissue culture dish for every three embryos prepared. Culture overnight in a 37 °C, 5% CO_2 incubator.

1.3.13) After the overnight incubation, replace the media with fresh 30 mL of media per plate. Place cells back into the incubator overnight.

Note! Check for Hcn4-GFP$^+$ cell contamination under a fluorescent microscope. The

culture should be GFP^-, and only become GFP^+ upon reprogramming.

1.3.14) The next day, harvest cells with pre-warmed fresh 3 mL of 0.25% trypsin-EDTA. Count and freeze cells. Typically, freeze cells at 3 x 10^6 cells per mL. The expected yield should be 3 x 10^6 cells per embryo.

2. Retrovirus production and reprogramming

Caution! The following protocol relies in the production and handling of infectious retroviruses. Perform the following steps in a Biosafety Level 2 cabinet under BSL-2 guidelines and sterile technique. Use 10% bleach to dispose of all materials exposed to retroviruses.

2.1) Retrovirus production and MEFs preparation

Note! The following protocol is adjusted to produce virus (in Plat-E cells) in a 6-cm plate format and the infection of MEFs in a 24-well plate. For other formats, refer to **Table 2.2.** MEFs will be plated at Day-1, so the timing will need to be coordinated appropriately for each experiment (Refer to section 2.3 and **Figure 2.2**).

2.1.1) Maintain Plat-E (PE) cells as per manufacturer's recommendations. Briefly, culture PE cells in DMEM supplemented with 10% FBS, 1 µg/mL puromycin, 10 µg/mL blasticidin, penicillin and streptomycin. Passage cells 1:4 every two days when the culture reaches 70-90% confluency.

2.1.2) **Day -2:** The day before transfection, seed Plat-E cells at 1 x 10^6 cells/well on a 6-well plate in transfection media. Plat-E cells should be 70-80% confluent at the time

of transfection.

2.2) Transfection using a commercial transfection agent.

Note! The commercial reagents (FuGene 6 and Opti-MEM) should be at room temperature (RT) before transfection. For the DNA transfection: Add each retroviral plasmid DNA individually (G, H, M, and T) to form a GHMT cocktail(Song et al., 2012b).

2.2.2) **Day -1:** In a 15 mL conical polystyrene tube, mix 60 µL of reduced serum media with 6 µL of transfection reagent per reaction for a 6-well plate format. Incubate the mixture for 5 min at room temperature.

Note! Since the transfection reagent used here binds to plastics, add directly to the reduced serum media to avoid any decrease in transfection efficiency.

2.2.3) Add a total of 2 µg of GHMT cocktail per reaction and gently tap to mix it. Do not vortex. Incubate the reaction for 15 min at RT.

2.2.4) Add the mixture from step 2.2.3 to the PE cells in a drop wise manner.

2.2.5) Incubate the transfected Plat-E cells overnight in a 37 °C, 5% CO_2 incubator. Record the time of transfection

2.3) Seeding of Hcn4-GFP mouse embryonic fibroblasts.

2.3.1) 1 h before plating MEFs, prepare the 24-well plate for immunocytochemistry.

2.3.1.1) Add a 12 mm fibronectin coverslip per well.

2.3.1.2) Coat wells with 300 µL of bovine collagen solution (e.g., SureCoat) and incubate in a 37 °C incubator for 1 h.

2.3.1.3) Aspirate coating solution immediately before MEF plating.

2.3.2) Thaw a frozen vial of Hcn4-GFP MEFs and wash x1 with pre-warmed fibroblast media at 500 x g for 5 min.

2.3.3) Determine cell viability using trypan blue exclusion or similar dyes. Calculate the number of viable cells per mL of culture use the following formula:

% viable cells = [1.00 - (Number of blue cells/ Number of total cells)] x 100

2.3.3.1) Calculate total number of viable cells use the following formula:

% viable cells x dilution factor x 10,000 x total volume of cell suspension

2.3.4) Seed 3 x 10^4 cells per well onto a 24-well plate with previously prepared bovine collagen solution-fibronectin coverslip.

2.4) Transduction and reprogramming of MEFs

Note! According to manufacturer's notes, properly maintained Plat-E cells produce an average titer of 1 x 10^7 infection units/mL. Although the titer is not directly measured for each experiment, a GFP control is used as a surrogate for infection efficiency of the viral batch. High GFP expression and intensity (GFP^+ >95%) correlates with successful GHMT-iCM.

2.4.2) **Day 0:** 24 h post-transfection, filter the PE retroviral medium through a 0.45 μm-pore size surfactant-free cellulose acetate filter and transfer to a 15 mL conical tube. Add polybrene to a final concentration of 8 μg/mL. Carefully replenish Plat-E dish with 2 mL of fresh transfection medium.

Note! Plat-E cells easily detach off the plate if media is changed too rapidly.

2.4.3) Aspirate the medium of the cultured MEFs and add the freshly collected retroviral

medium; it should yield 1.7 mL of media from a well of a 6-well plate. Add ~800 ul per well of a 24-well plate. Return MEF plate to the incubator and incubate overnight.

2.4.4) **Day 1:** Repeat steps 2.4.2 and 2.4.3. Discard Plat-E cells after the 2nd virus collection. Return induced MEFs to the incubator and let them rest overnight.

2.4.5) **Day 2:** 48 h post-induction, aspirate the Plat-E conditioned media and wash x1 with 1xPBS. Add 500 µL pre-warmed iCM media per well of a 24 well plate.

2.4.6) Replace iCM media every 2-3 days. Process plate 14 days after viral induction for immunocytochemistry (ICC) analysis of cardiac reprogramming.

3. Immunostaining of reprogrammed MEFs

3.1) 14 days post-induction, carefully aspirate the media.

3.2) Rinse each well with 300 µL of ice-cold 1x PBS. Aspirate excess solution.

3.3) Fix cells with 250 µL 4% paraformaldehyde (PFA) solution per well of a 24 well plate. Incubate 15 min at RT.

Note! Fixed cells can be stored in PBS at 4 °C for 1-2 weeks before staining.

3.4) Permeabilize cells by washing wells x3 with 300 µL 0.1% PBS-Triton X100 (PBST). Incubate 5 min at RT between washes. Aspirate excess solution after the last wash.

3.5) Block for 10 min at RT with 1x Universal Block Buffer at 300 µL/well.

3.6) Prepare (ICC) staining buffer: Add 1:1 of 1x PBS and 1x Universal blocking buffer. Dilute primary antibodies in ICC staining buffer and incubate antibodies overnight at 4 °C. Refer to the material section for recommended dilutions.

3.6.1) Stain one pair of slides with mouse α-actinin, chicken αGFP, and rabbit Nppa for iPM and iAM identification.

3.6.2) Stain one pair of slides with mouse α-actinin, chicken αGFP, and rabbit Myl2 for iPM and iVM identification.

3.7) The following day, wash wells x3 with 300 µL 0.1% PBST. Incubate 5 min at RT between washes. Aspirate excess solution after the last wash.

3.8) Prepare the secondary antibody dilutions in ICC staining buffer. Refer to the material section for recommended dilutions. Incubate secondary antibodies 1 h at RT, protected from light.

3.8.1) Stain all slides with the following secondary antibodies: mouse Alexa-555, chicken Alexa-488 and rabbit Alexa-647.

3.9) Wash wells x3 with 300 µL 0.1% PBST. Incubate 5 min at RT between washes. Protect from light.

3.10) Add 2.4 µL of antifade mounting media with 1.5 µg/mL of 4',6-diamidino-2-phenylindole (DAPI) to a glass microscope slide. Carefully remove the coverslip from the well of the 24-well plate, remove excess solution, and transfer to the glass slide with mounting media. Gently press the coverslip to remove excess volume and air.

3.11) Seal slides with preferred nail polish or plastic sealant. Store mounted slides at 4°C, protected from light (**Figure 2.5a**).

4. Identification of cardiac subtypes using confocal microscopy

Note! For imaging, a confocal microscope equipped with at least 2 fluorescent detectors capable of the spectral detection of 405, 488, 555 and 639 nm wavelengths is necessary to identify iPMs, iAMs, and iVMs. Image cells using a Plan-Apochromat 20x/0.75 or better.

Note! Using the manufacturer's image analysis software, scanning zoom images can achieve 40x-oil immersion quality images.

4.1) Image library: take 8-bit images with DAPI, Alexa-488, Alexa-555, and Alexa-647 channels (ch.). Pixel dwell time of 6 s, 1024 frame size, line step at 2, and averaging of 2 is sufficient for high resolution images.

4.2) For each slide, start from one edge and start scanning up and down in the red fluorescent channel (ch.) for α-actinin^{+} Sarcomere^{+} cells (Refer to **Figure 2.3** and **Figure 2.5A** for examples). Sarcomere striations are easier to identify visually in the 555 nm wavelength.

4.2.1) Once an α-actinin^{+} Sarcomere^{+} cell has been identified, switch to the green ch. and keep note if it is positive (iPM). Switch to the computer to assess the far-red 647 ch. (iAM or iVM).

Note! Cells that are α-actinin^{+} /Sarcomere^{+}/Hnc4-GFP^{+}/Nppa^{-}/Myl2^{-} are designated as iPMs. GFP expression will be seen throughout the cell (**Figure 2.4A**).

4.2.2) Stain slides with α-actinin (mouse-Alexa555), Hcn4-GFP (chicken-Alexa488), Nppa (rabbit-Alexa647), and DAPI. Cells that are α-actinin^{+}/Sarcomere^{+}/Hnc4-GFP^{-} /Nppa^{+} are iAM. Nppa staining will appear perinuclear and punctate (**Figure 2.4B**).

4.2.4) Stain slides with α-actinin (mouse-Alexa555), Hcn4-GFP (chicken-Alexa488), Myl2 (rabbit-Alexa647). Cells positive for α-actinin^{+}/Sarcomere^{+}/Hnc4-GFP^{-}/Myl2^{+} are iVMs. Myl2 staining will exhibit a striated form along the sarcomere filament. Due to variations in the quality of the staining and the Z-plane, striations may not be easily

visible (**Figure 2.4C**).

5. Quantification

Note! To assess the actual number of potential reprogrammable MEFs, 2 wells of a 24-well plate are seeded in parallel to the experimental wells and are harvested one day after plating. The total number of cells plated is then determined by averaging the two wells. This becomes the actual total cells plated (aTotal).

5.1) Sarcomere$^+$

5.1.1) Visually inspect each cell on a coverslip for proper α-actinin$^+$ /Sarcomere$^+$ (Right panels **Figure 2.3**) and record (**Figure 2.5B-i**).

5.1.2) Tabulate the total number of α-actinin$^+$ /Sarcomere$^+$ on each coverslip and divide by the actual total cells plated (aTotal) (**Figure 5b-iii**). For example, if aTotal = 12,500 cells, and 100 cells were α-Actinin$^+$/Sarcomere$^+$ then, 0.8% of the plated MEFs were reprogrammed. An average reprograming experiment will yield 1% α-Actinin$^+$ /Sarcomere$^+$ cells (**Figure 2.5C**).

5.2) Subtype$^+$

Note! For the following steps, refer to **Figure 2.5B-C** for a representative iCM quantification workflow. Briefly, for each sarcomere$^+$ cell, tabulate if it is unique for either subtype (**Figure 2.5B-i**). Calculate % Subtype (**Figure 2.5B-iii**) by dividing the number of subtype$^+$ cells over the average sarcomere$^+$ cell x 100 (**Figure 2.5B-i**). To calculate the

absolute % subtype efficiency (**Figure 2.5B-iv**), divide the subtype$^+$ cell number from **Figure 2.5B-i** by the total number of cells seeded x 100 (**Figure 2.5B-ii**).

5.2.1) For each of the α-actinin$^+$ /Sarcomere$^+$ cells, assess if they are either GFP$^+$/Nppa$^+$ or Myl2$^+$.

5.2.2) Tabulate total number of α-actinin$^+$/Sarcomere$^+$/Hnc4-GFP$^+$/Nppa$^-$/Myl2$^-$, α-actinin$^+$/Sarcomere$^+$/Hnc4-GFP$^-$/Nppa$^+$, and α-actinin$^+$/Sarcomere$^+$/Hnc4-GFP$^-$/Myl2$^+$ (**Figure 2.5B-i**).

5.2.3) To calculate the ratio each % Subtype, divide the total Subtype$^+$ cell over the total number of α-actinin$^+$/Sarcomere$^+$ in that well and multiply by 100. GHMT generates iPMs, iAMs and iVMs at equal ratios (**Figure 2.5B-iii**).

5.2.4) To calculate the absolute subtype$^+$, divide the total number of subtype$^+$ for the experimental condition over aTotal and multiply by 100 (**Figure 2.5B-ii**). On average iPMs represent 0.3% of the total reprogrammed population, iAMs 0.3%, and iVMs 0.25% (**Figure 2.5B-iv**).

REPRESENTATIVE RESULTS

Taking advantage of the PM-specific reporter mouse, it was possible to develop a multiplex immunostaining strategy to identify the diverse endogenous myocytes, as depicted in **Figure 2.1**. Following the reprogramming steps shown in **Figure 2.2**, induction of subtype-specific CMs can be detected as early as day 4(Nam et al., 2014b), albeit, at a low-rate. By day 14, the experiment can be stopped and assessed for sarcomere

organization (**Figure 2.3**), and subtype-specification (**Figure 2.4**). **Figure 2.5** summarizes the workflow for the slide preparation for ICC (**Figure 2.5A**), and the quantification of iCM subtype-specific cells **(Figure 2.5B-C**).

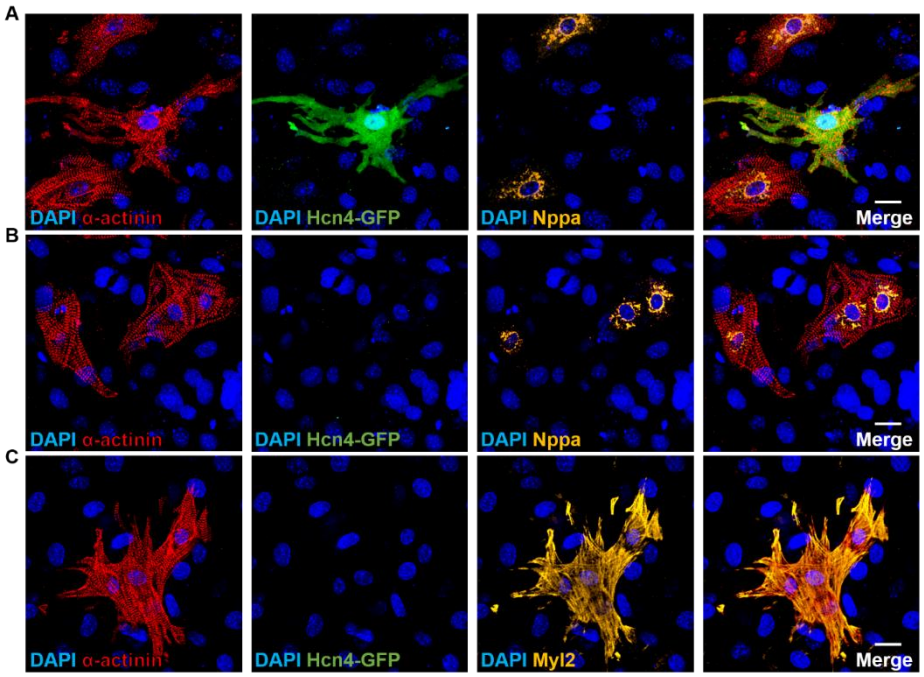


Figure 2.1. Subtype diversity of endogenous cardiomyocytes. A-B) *Immunocytochemistry (ICC) staining of neonatal atrial cardiomyocytes from Hcn4-GFP reporter mice for α-actinin (sarcomere marker, red), Hcn4-GFP (PM marker, green), and Nppa (atrial marker, orange).* ***C)*** *Immunocytochemistry staining of neonatal ventricular cardiomyocytes from Hcn4-GFP reporter mice for α-actinin (sarcomere marker, red), Hcn4-GFP (PM marker, green), and Myl2 (ventricular marker, orange). DAPI (blue): nuclear staining. Scale bars: 20 μm.*

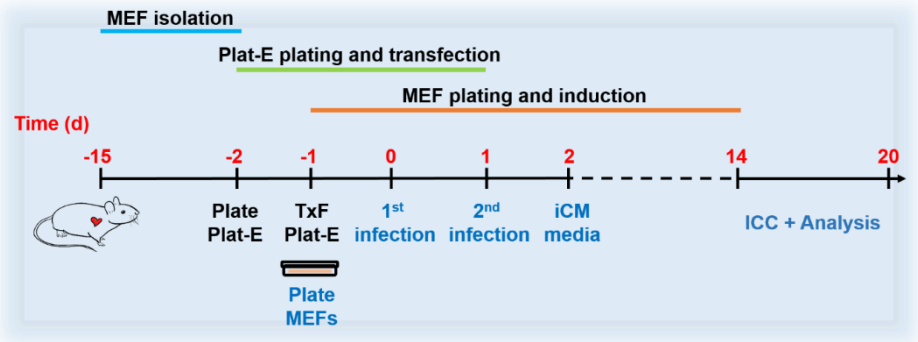


Figure 2.2. Induced reprogramming schematic timeline. *Schematic representation of the GHMT-induced Hcn4-GFP MEFs. The three major stages are depicted.*

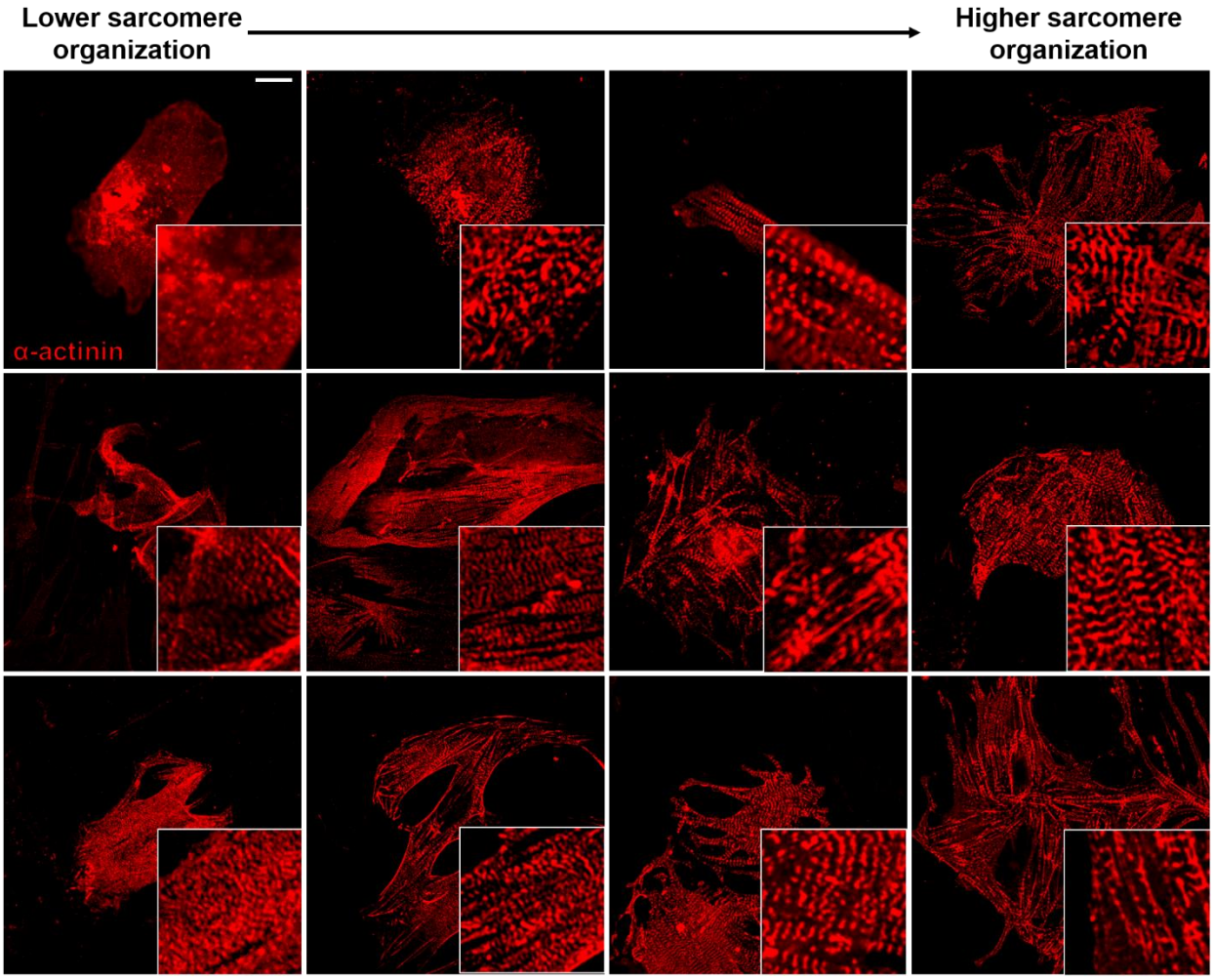


Figure 2.3. Degree of sarcomere organization. *ICC staining of Hcn4-GFP MEFs 14 days after GHMT transduction for α-actinin (sarcomere marker, red) shows a diverse range of sarcomere organization. The degree of organization increases from left to right panels. Representative pictures of each level (n= 3). Scale bars: 20 μm.*

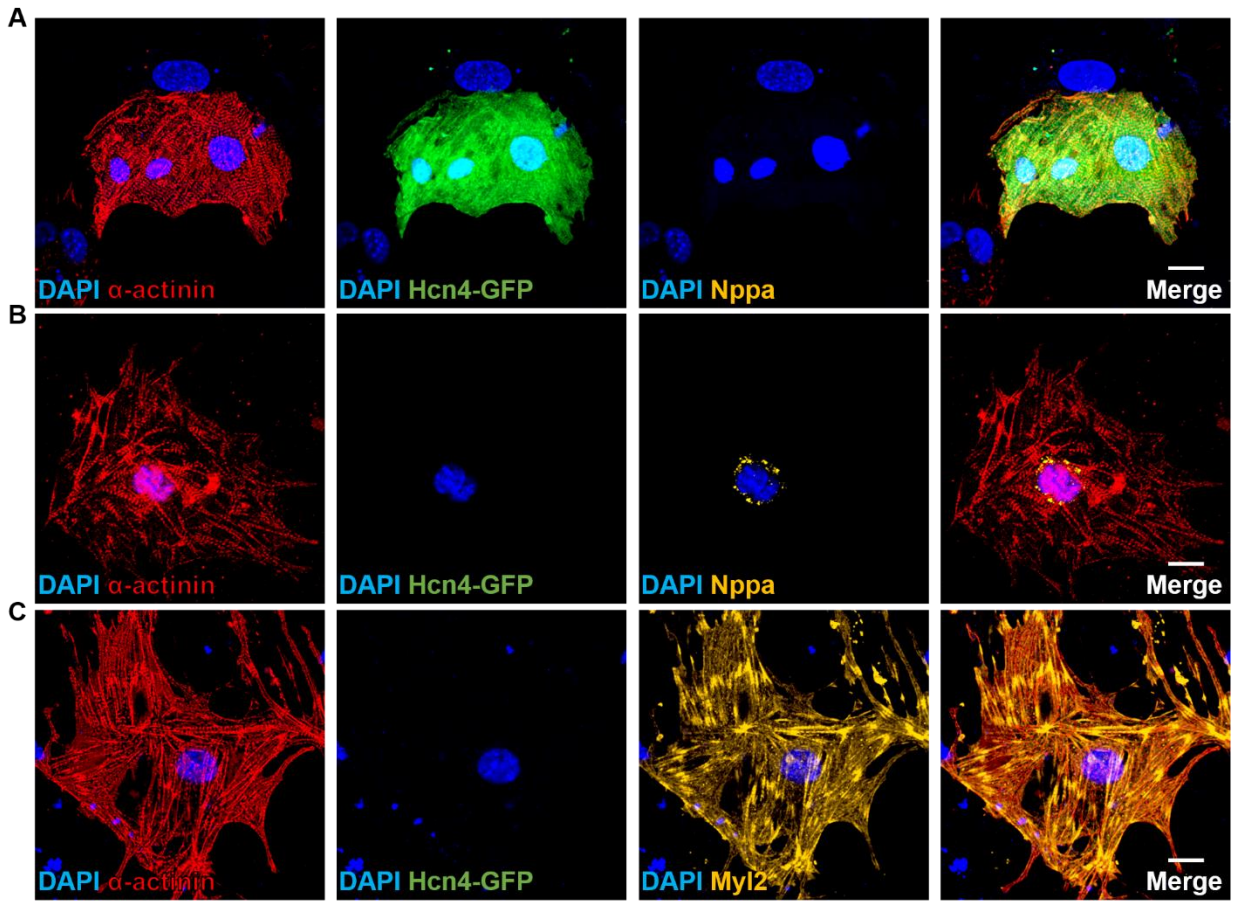


Figure 2.4. Subtype-specific reprogrammed cardiomyocytes. A-C) *ICC staining of GHMT-transduced Hcn4-GFP MEFs for α-actinin (sarcomere marker, red), Hcn4-GFP (PM marker, green), Nppa (atrial marker, orange) or Myl2 (ventricular marker, orange). DAPI (blue): nuclear staining. Scale bars: 20 μm.*

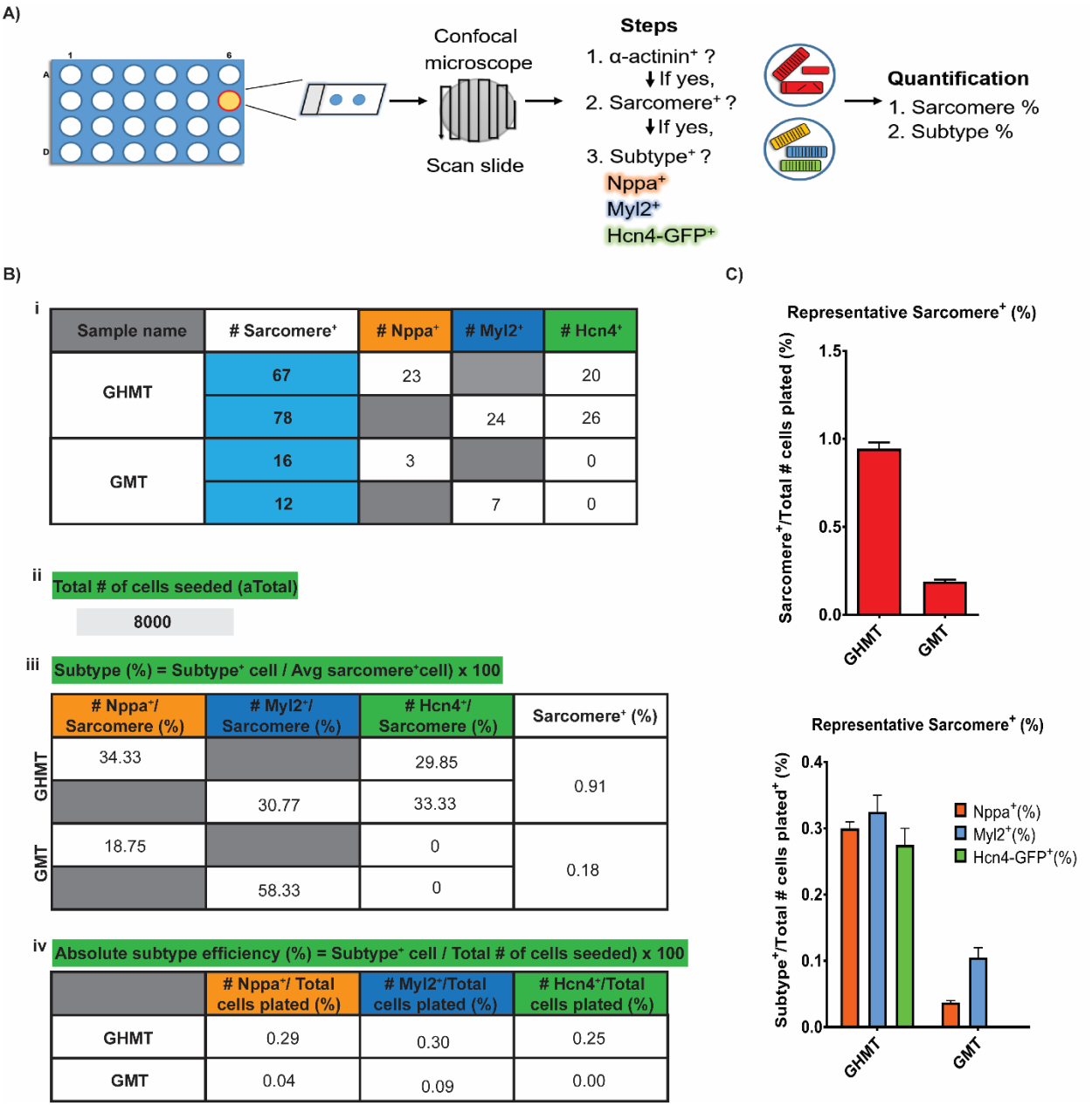


Figure 2.5. Image acquisition and analysis workflow. *Schematic representation for the image analysis.* ***A)*** *depicts the priority order to assign sarcomere+ and subtype-specificity to a cell.* ***B) and C)*** *show the expected results from an average GHMT- iCM experiment. Key points and formulas are shown in green.*

iCLM media		
Component	**Volume (mL)**	**Final concentration**
DMEM	270	
Medium 199	90	
FBS	50	10%
Insulin-Transferrin-Selenium G	2.5	0.50%
MEM vitamin solution	10	2%
MEM Amino Acids	20	4%
Non-essential amino acids	10	2%
Antibiotic-Antimycotics	10	2%
B-27 supplement	10	2%
Heat inactivated Horse Serum	25	5%
NaPyruvate	2.5	1.5mM
Plat-E media (PE)		
Component	**Volume (mL)**	**Final concentration**
DMEM	450	
FBS	50	10%
Penicillin/Streptomycin	5	1%
Puromycin	0.05	1ug/mL
Blasticidin	0.5	10ug/mL
Fibroblast medium (FB)		
Component	**Volume (mL)**	**Final concentration**
DMEM	450	
FBS	50	10%
Penicillin/Streptomycin	5	1%
Glutamax	5	1%
Transfection medium (TxF)- Filtered (0.45um)		
Component	**Volume (mL)**	**Final concentration**
DMEM	450	
FBS	50	10%
Immunocytochemistry (ICC) staining buffer		
Component	**Volume (mL)**	**Final concentration**
1x PBS	5	
1X Universal blocking buffer	5	

Table 2.1. Culture medium. *Table summary for the preparation of the several mediums used during GHMT-induced reprogramming.*

A) Plat-E seeding and transfection

Plate/Dish	Surface Area (cm^2)	Seeding density	Growth medium (mL)	Total DNA amount to transfect (ug)	FuGene 6 (uL)	Opti-MEM (uL)
15cm plate	152	10e6 cells	20	25	75	600
10cm plate	55	5.5e6 cells	10	9	27	300
6cm plate	21	2.2e6 cells	4	3.5	10.5	105
6well/x1	9	1e6 cells	2	2	6	60
12well/x1	4	4.0e5 cells	1	0.5	1.5	15
24well/x1	2	2.0e5 cells	0.5	0.3	0.9	9
48well/x1	1	1.7e5 cells	0.25	0.15	0.45	4.5

B) Fibroblast seeding and induction

Plate/Dish	Fibroblast seeding density (millions)	Approximate infection units for iCLM
6cm plate	0.22-0.33	5x10^7 (~5mL)
6well/x1	0.1-0.15	3x10^7 (~3mL)
12well/x1	0.04-0.06	1.3x10^7 (~1mL)
24well/x1	0.02-0.03	0.65x10^7 (~0.8 mL)
48-well	0.001-0.015	0.3x10^7 (~0.4 mL)

***Table 2.2. Seeding, transfection and induction formats**. **A)** Table summary for the plating and transfection of Plat-E cells. **B)** Seeding density and approximate infection units (or viral supernatant) needed to induce MEFs into cardiomyocyte-like cells.*

DISCUSSION

The present study provides a direct-reprogramming strategy for the conversion of MEFs into a diverse set of cardiac subtypes via the retrovirus-mediated expression of the cardiac transcription factors Gata4, Mef2c, Tbx5, and Hand2 (GHMT). Using a multiplex immunostaining approach in combination with a PM-specific reporter mouse, we can identify iAM, iVMs and iPMs at single cell resolution. Such an assay allows for an experimental *in vitro* system capable of isolating the contributions of individual transcription factors towards subtype diversity and sarcomere development. In parallel, this could bring insight to new transcription factors or small molecules that bias iCMs into a specific lineage. Nevertheless, there are several critical steps for the successful completion of this assay. Below, we address the impact of viral titer, fibroblast quality, and imaging analysis in a general iCM experiment.

In our study we employ ecotropic-retroviruses to reprogram E12.5 MEFs. We noticed the retroviral titer has a direct relationship with the quality of the Plat-E cells. High passage number (>35) and poor culturing techniques severely affect the quality of the retroviral particles; therefore, there are several considerations to keep in mind. Plat-E cells do not produce VSV-G pseudotyped virus, and are thus, unable to withstand ultracentrifugation or freezing cycles(Burns et al., 1993; Ichim and Wells, 2011). To preserve the longevity of the Plat-E it is imperative to maintain the stock with antibiotic selection. However, they should be maintained in antibiotic free media upon viral production. In our experience, the transfection reagent used here provides the highest transfection efficiencies in Plat-E cells. If other transfection methods are to be used, comparing the viral titers produced is essential(Qian et al., 2013). Although there are recommendations by the manufacturer to harvest the viral supernatant 48 h after transfection, we observed that two 24-h harvesting rounds yield higher reprogramming efficiencies while avoiding toxic effects usually associated with higher-titer viral preps. Furthermore, though several studies have shown the feasibility of commercial viral supernatant concentrators(Yang et al., 2014a), we have not employed these in our regular protocol in order to maintain a higher throughput.

In addition to high titer viral cocktails, fibroblast quality is of crucial importance for a successful reprogramming assay(Muraoka and Ieda, 2014). If timed correctly, freshly isolated MEFs should be utilized due to their higher efficiencies compared to frozen stocks. This could be related to the nature of retroviruses, as they need a highly-proliferative host in order to integrate(Coffin et al., 1997). Additionally, MEF seeding density plays a critical role. We have included a table with the seeding densities employed

in our experiments **(Table 2.2).** Moreover, passaging the MEFs will also significantly decrease reprogramming efficiency.

Immunocytochemistry (ICC) is our standard technique for analysis of sarcomere organization and subtype specification. With the help of a PM-GFP reporter mouse, we were able to form an antibody panel for the detection of three major cardiac subtypes (AM, VM, and PM). However, due to constraints of antibody species availability and the limitation of 4-channels on a standard confocal microscope set-up, two coverslips per subtype are needed to quantify the prevalence of all three subtypes. One coverslip will stain for α -Actinin/GFP(Hcn4)/Myl2, and one for α -Actinin/GFP(Hcn4)/Nppa. Based on our previous observation that sarcomeric structure is a common characteristic of all CMs and a potential prerequisite for subtype specification(Nam et al., 2014b), the first step in our analysis is determining sarcomere$^+$ cells. Yet, due to its subjective nature, establishing the level of sarcomere organization is perhaps the most difficult part of this assay; this can be limited by averaging multiple observer's quantifications or by developing computational cell segmentation software to automate the process(Bass et al., 2012a). Using endogenous cells as a point of reference, we discovered a threshold for well-organized sarcomere$^+$ and utilized that to score iCLMs **(Figure 2.3).** Given these parameters, an average experiment will give rise to 20-30% α -Actinin$^+$ cells, while only 1% α -Actinin$^+$/Sarcomere$^+$. Of the 1% sarcomere$^+$, ~30% will be Nppa$^+$, Myl2$^+$, or Hcn4-GFP$^+$.

Cardiomyocytes are structurally complex systems and using mRNA or flow cytometry analysis are not enough to capture the morphological changes during iCLM reprogramming. However, due to the specialized skills and tools required for more

stringent analysis, like patch clamping or Ca^{2+} imaging, these types of studies may be not readily available to many labs. Thus, the described methodology is unique in that it provides an almost general approach to study key parameters of iCM development with a compromise between functionality and throughput studies.

Despite the many recent advances in direct reprogramming, much work remains to be done to better understand the molecular mechanism that regulates cardiac reprogramming, and more specifically, subtype specification. These mechanisms will become especially important to translate direct reprogramming for clinical applications. As such, in this study we describe a platform capable of directly modulating discrete parameters to assess the contribution towards sarcomere development, subtype specification, and iCM maturity. Moreover, this system can be further developed to work in a high-throughput format allowing for complex screening of small molecules or extracellular matrixes for the next step in regenerative cardiology.

CHAPTER 2 ACKNOWLEDGEMENTS

A.F.-P. was supported by the National Science Foundation Graduate Research Fellowship under Grant No.2015165336. N.V.M was supported by grants from the NIH (HL094699), Burroughs Wellcome Fund (1009838), and the March of Dimes (#5-FY14-203). We acknowledge Young-Jae Nam, Christina Lubczyk, and Minoti Bhakta for their important contributions to protocol development and data analysis.

CHAPTER THREE

AUTOMATING SUBTYPE QUANTIFICATION

The contents of this chapter are reproduced from:

> Sutcliffe, M. D., Tan. P. M**., Fernandez-Perez, A.**, Nam, YJ., Munshi, N. V., Saucerman, J. J. High content analysis identifies unique morphological features of reprogrammed cardiomyocytes. Scientific Reports. DOI: 10.1038/s41598-018-19539-z (2018).

Chapter Overview

The previous chapter illustrates the manual methodology that I used to generate the majority of my iCM data. Although the quantification yields a high degree of accuracy in determining cardiac subtypes, the system is far from been used in an unbiased matter and high-throughput projects. Keeping this in mind, we envisioned an automated subtype quantification system with a long-term goal of being user friendly and readily available to the scientific community. However, given the skill set required to generate such automatization, and our lack of expertise in the matter, we collaborated with the Sarucerman lab at the University of Virginia to develop SarcOmere Texture Analysis (SOTA), an improved Haralick texture package for the CM-specific study of spatial grey tone values. Given the importance of this tool to projects like mine, I have incorporated our collaboration in this thesis. I highly recommend using this tool for any staining-driven study. My contribution to this publication resides in the collection of all the confocal images from endogenous and reprogrammed cells to generate the training sets necessary for this algorithm to work.

ABSTRACT

Direct reprogramming of fibroblasts into cardiomyocytes is a promising approach for cardiac regeneration but still faces challenges in efficiently generating mature cardiomyocytes. Systematic optimization of reprogramming protocols requires scalable, objective methods to assess cellular phenotype beyond what is captured by transcriptional signatures alone. To address this question, we automatically segmented reprogrammed cardiomyocytes from immunofluorescence images and analyzed cell morphology. We also introduce a method to quantify sarcomere structure using Haralick texture features, called SarcOmere Texture Analysis (SOTA). We show that induced cardiac-like myocytes (iCLMs) are highly variable in expression of cardiomyocyte markers, producing subtypes that are not typically seen in vivo. Compared to neonatal mouse cardiomyocytes, iCLMs have more variable cell size and shape, have less organized sarcomere structure, and demonstrate reduced sarcomere length. Taken together, these results indicate that traditional methods of assessing cardiomyocyte reprogramming by quantifying induction of cardiomyocyte marker proteins may not be sufficient to predict functionality. The automated image analysis methods described in this study may enable more systematic approaches for improving reprogramming techniques above and beyond existing algorithms that rely heavily on transcriptome profiling.

INTRODUCTION

Cardiomyocytes have limited regenerative capacity in the adult heart, and following a myocardial infarction many cardiomyocytes are irreversibly lost(Tzahor and Poss,

2017). In response, activated fibroblasts proliferate, migrate into the injured area, and deposit collagen and other extracellular matrix proteins(Ma et al., 2017; Prabhu and Frangogiannis, 2016). A scar forms and, over time, the contractile function of the heart weakens, leading to congestive heart failure. To contend with this growing clinical problem, methods for generating new cardiomyocytes are greatly needed. For example, pluripotent stem cells (i.e. ESCs or iPSCs) can be expanded and differentiated ex vivo prior to transplantation. Although this approach has shown promise in large animal models(Chong et al., 2014; Shiba et al., 2012) and tumors have not been observed to date, the use of pluripotent stem cells raises concerns of teratogenicity. An alternative therapeutic strategy that bypasses the concerns of the cell transplantation approach involves direct conversion of fibroblasts into functional cardiomyocytes(Sadahiro et al., 2015; Srivastava and DeWitt, 2016b). This approach is particularly attractive, since it can be accomplished in activated fibroblasts in situ to convert them into cardiomyocytes rather than form scar tissue(Qian et al., 2012a; Song et al., 2012a).

Direct reprogramming involves transduction of various combinations of transcription factors that typically consist of key developmental regulators. The most commonly described transcription factor combination for induced cardiac-like myocyte (iCLM) reprogramming includes Gata4 (G), Mef2c (M), and Tbx5 (T), with or without Hand2 (H)(Ieda et al., 2010a; Song et al., 2012a). GMT and GHMT both convert fibroblasts into functional cardiomyocytes in vitro and in vivo(Inagawa et al., 2012; Qian et al., 2012a). Interestingly, in vivo reprogramming is substantially more efficient than in vitro reprogramming, suggesting that various aspects of the endogenous milieu are likely to influence reprogramming efficacy. Nevertheless, the precise mechanisms responsible

for reprogramming, and the ideal transcription factor combinations required to produce mature subtype-specific cardiomyocytes remain unclear. One major hurdle that has slowed progress in this field is the lack of objective and quantitative measures of cardiomyocyte reprogramming. We recently found that GHMT generates all three cardiomyocyte subtypes (i.e. atrial, ventricular, and pacemaker) but with low efficiency due in part to incomplete sarcomere formation(Nam et al., 2014a). Thus, we sought to develop an unbiased algorithm for assessing cardiomyocyte subtype and sarcomere structure in directly reprogrammed fibroblasts.

Automated image processing algorithms to extract morphological and textural information provide objective and quantitative methods to analyze cells. These methods may be used to assess the function of induced cardiomyocytes. Cardiomyocytes are composed of bundles of myofibrils, each of which consists of distinct, repeating sarcomeres. Thus, the sarcomere is the basic force-generating unit of striated muscle. Sarcomeres are composed of myosin and actin, the two components of cross-bridge formation, and Z-lines, which are protein complexes defining the edges of the sarcomeres. It is intuitive that clearer sarcomere structure, as indicated by Z-line structure in the α-actinin stain, is correlated with cardiomyocyte functionality. We have previously shown that sarcomere organization is a prerequisite for reprogrammed cardiomyocytes to spontaneously contract, a well-established parameter of functionality(Nam et al., 2014a). To usefully reprogram cardiomyocytes, therefore, careful attention must be given to both cell morphology and contractility.

To quantify and thus compare iCLMs with neonatal mouse cardiomyocytes, we have developed a fully auto- mated method of segmenting cells from multi-channel

immunofluorescence images and used this to analyze their morphology. To quantify sarcomere structure, we developed a method, based on offset distance–angle distributions of Haralick texture features, called SarcOmere Texture Analysis (SOTA). Using this method, we found that current methods of direct reprogramming generate cardiomyocytes with less organized sarcomeres, shorter sarcomere lengths, as well as an apparent lack of coordination between cellular elongation and sarcomere alignment. These new automated image analysis methods may facilitate quantitative screening of experimental protocols that further enhance the efficiency and fidelity of cardiomyocyte reprogramming.

RESULTS

Automated subtype classification and morphological analysis of induced cardiac-like myocytes. We previously developed algorithms for automated cell segmentation and morphological analysis of primary neonatal cardiomyocytes based on a combination of DAPI and α-actinin(Bass et al., 2012b). Here, we extended that method to include multiple cardiomyocyte markers (α-actinin, Hcn4, and Nppa) that distinguish the diverse cell subtypes that arise from cell reprogramming by GHMT transduction. Hcn4 is an ion channel predominantly expressed in pacemaker cardiomyocytes and Nppa is a perinuclear marker of atrial cardiomyocytes(Nam et al., 2014a). The heterogeneous expression of these three markers in reprogrammed cardiomyocytes required a more flexible method for systematically identifying and classifying nuclei, cell borders, and cell subtypes.

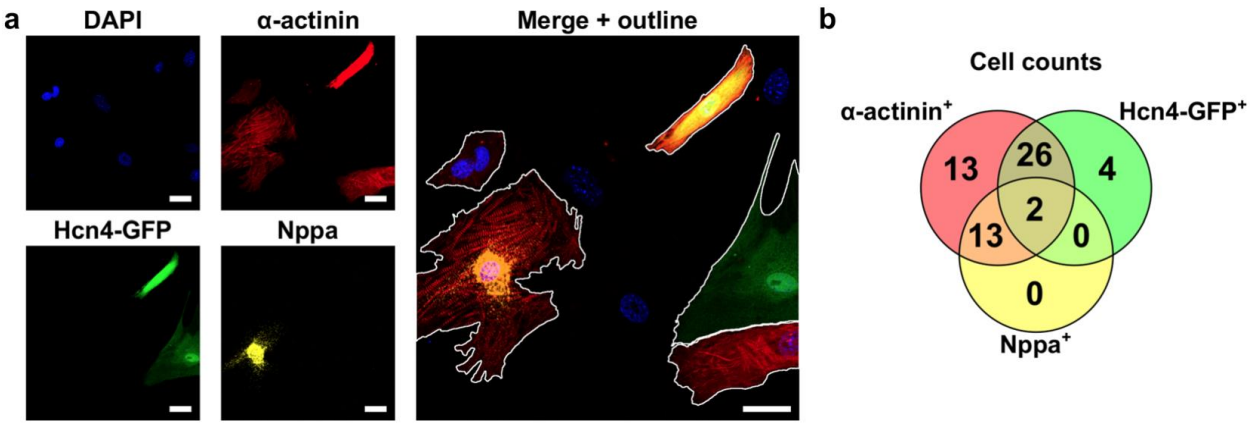


***Figure 3.1**. Automated cell segmentation identifies induced cardiac-like myocytes (iCLMs) with heterogeneous marker expression. **A)** Mouse embryonic fibroblasts were isolated from E13.5 mouse embryos and reprogrammed by transfecting with retroviral constructs of GHMT. Cells were examined by immunofluorescence for expression of cardiac markers α-actinin, Hcn4-GFP, and Nppa. The merged image is shown on right with automated cell segmentation outlines in white. Scale bar = 20 μm. **B)** Venn diagram showing the distribution of cardiac marker expression among iCLMs. Only iCLMs that expressed at least one marker are shown. Most cells (132) were not positive for any cardiac marker.*

Neonatal mouse atrial cardiomyocytes (CMs) and GHMT-reprogrammed cells were fixed and stained for DAPI, α-actinin, Hcn4-GFP, and Nppa. Nuclei were first identified by thresholding the processed DAPI channel using Otsu's method(Otsu and cybernetics, 1979). The α-actinin and Hcn4-GFP images were similarly thresholded by Otsu's method. Nuclei that were fully within α-actinin+ or Hcn4-GFP$^+$ regions were classified as positive for those respective cardio- myocyte markers. Nppa classification was based on the 90th percentile intensity within the perinuclear region, defined as the area extending 8 pixels (1.25 μm) from the nuclear boundary. Nuclei of the same classification group that were within 25 pixels (3.91 μm) of one another were joined and assumed to be part of a binucleated cell, as validated previously(Bass et al., 2012b).

Cell boundaries were segmented based on sequentially masked images to better distinguish between neighboring cells with distinct α-actinin$^+$ and Hcn4-GFP$^+$ expression.

First, Hcn4-GFP$^+$ regions were masked to allow segmentation of α-actinin+/Hcn4-GFP− cells. Next, α-actinin+ regions were masked to segment α-actinin−/ Hcn4-GFP$^+$ cells. Finally, the inverse of the intersection of these two regions was used to mask a merged α-actinin/Hcn4-GFP image to segment α-actinin$^+$/Hcn4$^+$ cells. In all cases, segmentation was performed via the watershed method of the gradient-transformed image, which finds regions of maximally changing intensities(Vincent et al., 1991). An example segmented image is shown in **Figure 3.1**.

Most GHMT-transduced cells were negative for all cardiomyocyte markers, indicating the low efficiency of current reprogramming methods(Nam et al., 2014a). GHMT-transduced cells were classified as induced cardiac-like myocytes (iCLM) if they expressed α-actinin. Further confirmation of cardiomyocyte induction and maturity was assessed by pericentriolar material 1 (PCM1) staining{Zebrowski, 2015 #698} (**Supplementary Figure S1**). The GHMT-transduced cells expressed five out of seven possible combinations of the α-actinin, Hcn4-GFP, and Nppa markers (**Figure 3.1B**). Induced atrial-like cells (α-actinin$^+$/Nppa$^+$) and induced pacemaker-like cells (α-actinin$^+$/Hcn4-GFP$^+$) were identified, suggesting that these methods can produce cardiomyocytes similar to the defined phenotypes found in vivo. Furthermore, we have previously shown that induced atrial-like cells also express Myl7, an additional atrial cardiomyocyte marker(Nam et al., 2014a). In addition, several unexpected combinations were found. Some cells were only positive for the Hcn4-GFP marker and could be incompletely reprogrammed or mis-programmed cells. Although Hcn4 is expressed widely during heart development(Liang et al., 2013b), we have previously shown that GHMT-transduced fibroblasts do not pass through an Nkx2.5 lineage-positive

intermediate(Nam et al., 2014a). Therefore, it is more likely that α-actinin^{-}/ Hcn4-GFP^{+} cells represent mis-specification of cell state, as Hcn4 is also expressed in specific regions of the central nervous system, including the cerebellum(Zúñiga et al., 2016). Additionally, two cells were identified and manually con- firmed to be α-actinin^{+}/Hcn4-GFP^{+}/Nppa^{+}. These cells tended to have more visually distinguishable sarcomeres and may represent so-called transitional cells that surround the sinoatrial node(Abad et al., 2017b). Finally, several cells were only α-actinin^{+}, which may be induced ventricular-like myocytes(Nam et al., 2014a). While this GHMT method of reprogramming has been shown to generate α-actinin^{+}/Myl2^{+} cells, we could not confirm this within our current limitation of visualizing four separate fluorescent channels (DAPI, Hcn4-GFP, Nppa, and α-actinin). However, we have previously shown that Nppa^{+} reprogrammed cells do not express Myl2(Nam et al., 2014a).

After automated segmentation, cells were analyzed for area and other morphological characteristics (**Figure 3.2**). iCLMs and CMs had similar median cell areas, however iCLM area was more variable. Indeed, several iCLMs were substantially larger than the range seen for endogenous CMs. We found that most of these larger iCLMs were α-actinin^{+}/Hcn4-GFP^{+} (**Supplementary Figure S3.2**). Higher variability of iCLMs was also seen in cell circularity. In contrast, eccentricity and elongation, which are related to overall aspect ratios, were similar between iCLMs and CMs, with most cells exhibiting a major/minor axis ratio of about 2:1. This morphological variability suggests that future reprogramming techniques could benefit from improved methods for controlling cardiomyocyte size.

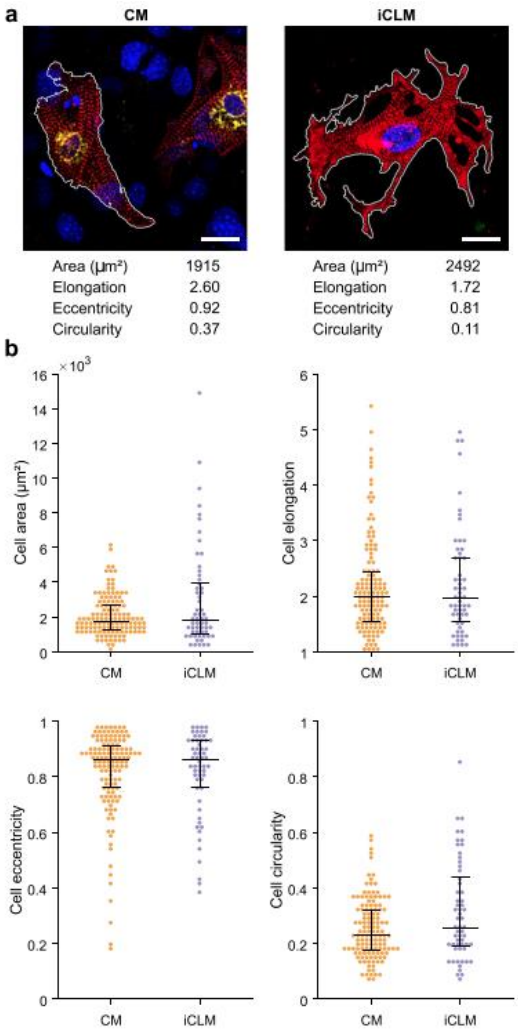


Figure 3.2. Increased morphological variability of reprogrammed iCLMs compared to endogenous CMs. A) *Segmented immunofluorescence images of endogenous CMs (left) and reprogrammed iCLMs (right), with automated morphology measurements for these cells below. Scale bar = 20 µm.* ***B)*** *Population measurements of various morphological features. Black bars represent 25th, 50th, and 75th percentile.*

Quantitative metrics for measuring sarcomere organization. The presence of visually distinguishable sarcomeres is frequently used to indicate cardiomyocyte maturity and functionality(Abad et al., 2017b; Liang et al., 2013b). Most current methods involve subjectively analyzing sarcomere structure or using Fourier transforms on manually selected rectangular sub-regions to quantify sarcomere organization(Pasqualini et al.,

2015), potentially introducing bias. To address these limitations, we developed a pixel-based image analysis method for assessing sarcomere structure, called SarcOmere Texture Analysis (SOTA). SOTA utilizes Haralick texture features, which are calculated from the gray level co-occurrence matrix (see Methods), that can be applied to any geometric shape(Haralick et al., 1973; Haralick and Shapiro, 1992). In SOTA, one of 13 Haralick texture features is computed for a range of orientations (0° to 180°) and pixel offset distances, forming an offset distance– angle distribution from which various features of sarcomere structure can be extracted.

We first applied SOTA to images with stripes that are representative of idealized sarcomeres (**Figure 3.3).** As shown in **Figure 3.3A-B**, such stripes produced repeated peaks in Haralick correlation, with the greatest magnitude in the direction of the sarcomeres. Increasing sarcomere length spreads these correlation peaks (**Figure 3.3C**), while introducing noise reduces the magnitude of the correlation peaks (**Figure. 3.3D**). The rate of decay in correlation peaks in the sarcomere direction is sensitive to the persistence of serially aligned sarcomeres (**Figure 3.3),** while decay in the longitudinal direction is sensitive to the width of the sarcomere bands (**Figure 3.F**).

We then applied SOTA to representative neonatal mouse CMs that had been subjectively classified as having highly organized or disorganized sarcomeres (**Figure 3.4**).

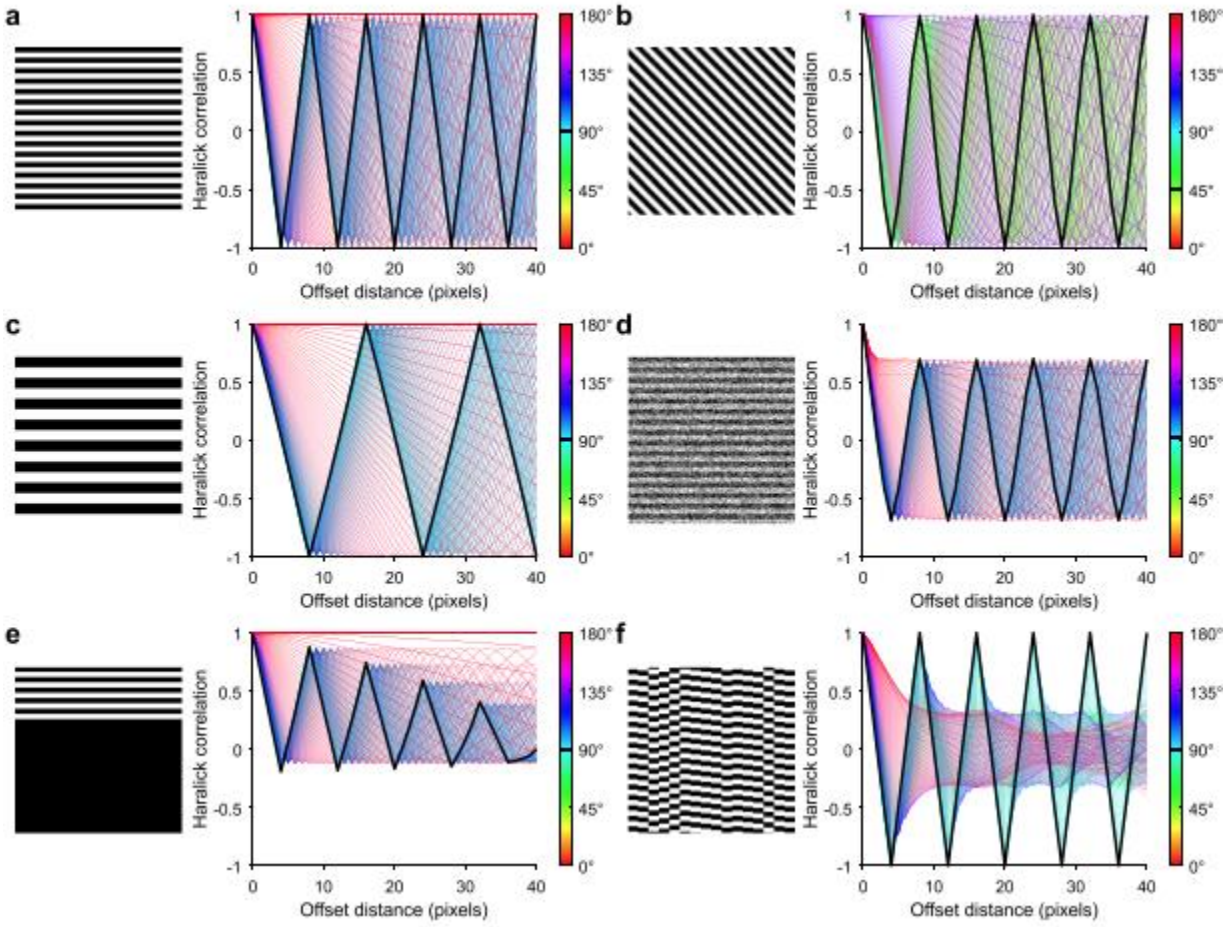


Figure 3.3. Haralick correlation metric in idealized images. A) *Horizontal stripes.* ***B)*** *Diagonal stripes.* ***C)*** *Horizontal stripes of a different frequency.* ***D)*** *Horizontal stripes with noise added.* ***E)*** *Horizontal stripes in one third of the image.* ***F)*** *Bands with random vertical offset.*

In cells with organized sarcomeres, a characteristic striated pattern is observed in the α-actinin stain. The corresponding Haralick offset distance–angle distributions of real cells decay more rapidly towards zero with increasing offset distance, because such pixels are less likely to be of similar intensity. At the angle corresponding to the fiber direction, this trace develops a decaying sinusoidal pattern, with peaks representing sets of pixels of similar intensity a specified distance apart (**Figure 3.4B**). Biologically, these peaks are indicative of adjacent Z-lines within the cytoskeletal structure. Sarcomere

organization is quantified by calculating the maximum peak prominence of all the traces. The angle at which this maximum occurs is the primary sarcomere direction. Sarcomere length is the pixel offset distance of the maximum peak prominence.

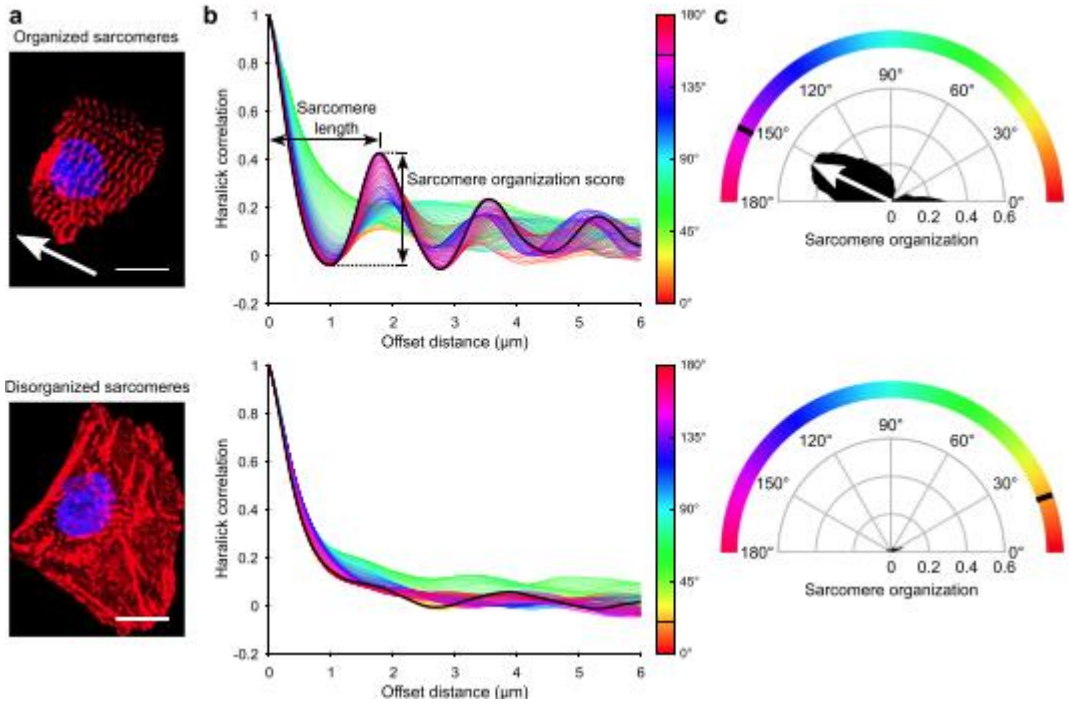


Figure 3.4. Automated measures of sarcomere organization and sarcomere length. ***A)*** *Masked immunofluorescence images of neonatal mouse CMs with organized (top) or disorganized (bottom) sarcomeres. Scale bar = 10 μm.* ***B)*** *Haralick correlation metric computed at multiple offset distances and angles to determine sarcomere organization and sarcomere length. Sarcomere organization score is the maximum amplitude of the decaying sinusoidal trace. Sarcomere length is the distance to the first peak.* ***C)*** *Sarcomere organization assessed as a function of angle to assess the primary direction of sarcomere alignment. Arrow points in the direction of sarcomere alignment and is repeated in* ***A)****. Color bar in middle panel aligns with circumferential color bar in right panel.*

To assess the performance of SOTA as well as previously proposed metrics of sarcomere organization(Bass et al., 2012b; Pasqualini et al., 2015; Sheehy et al., 2014) sets of neonatal mouse atrial or ventricular CMs with highly organized (n = 32) or disorganized (n = 26) sarcomeres were compared. Multiple variations of SOTA using

different Haralick texture features were used, in addition to Gabor filters and Fourier transforms (**Figure 3.5**). Gabor filters use sinusoidal plane waves at specified orientations multiplied by a Gaussian function to detect edges in images(Fogel and Sagi, 1989). These filters were applied at multiple orientations and wavelengths, and the maximum periodic response magnitude was used as the metric. Fourier transforms are also used to convert an image to the frequency domain to assess the repeating sarcomere structure(Pasqualini et al., 2015; Sheehy et al., 2014).

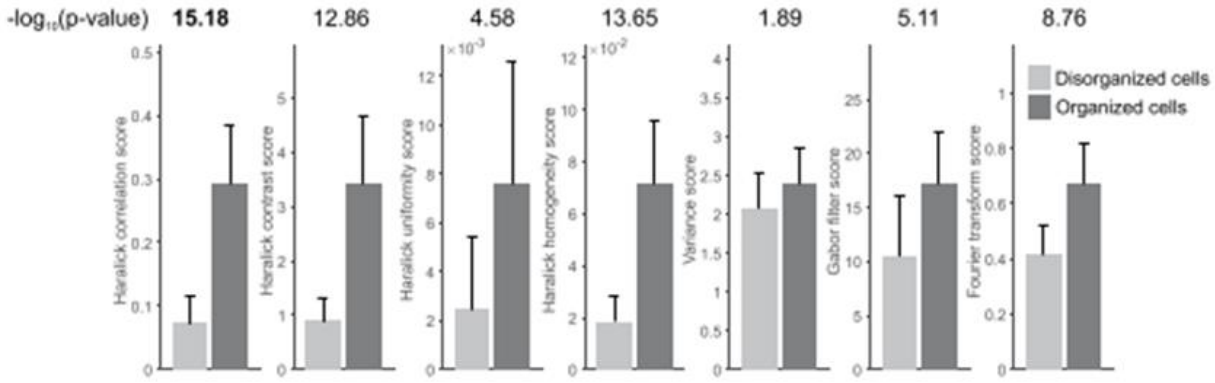


Figure 3.5. Comparison of sarcomere organization metrics in discriminating organized and disorganized cardiomyocytes. *Selected highly organized (n = 32) and disorganized (n = 26) neonatal mouse cardiomyocytes were used to quantitatively compare methods for measuring sarcomere organization. Out of 13 Haralick pixel-based measurements, 4 were identified as candidates for sarcomere organization measurements. −log10(p-values) are reported above each pair. P-values are calculated by two-sample t-tests. Error bars show standard deviation.*

Among these various methods, the SOTA method based on the Haralick correlation metric best distinguished between cells with highly organized and disorganized sarcomeres. This may be in part due to the method's ability to analyze pixels within the cellular region alone. In comparison, Fourier transforms must be applied to either the bounding box of an image, or a sub-region of the cell. Analyzing the bounding

box image introduces artifacts associated with cell shape, while selecting a sub-region of the cell would leave out information or introduce bias and be less amenable to automation.

Most reprogrammed cardiomyocytes have lower sarcomere organization. We next used SOTA to compare sarcomere organization in reprogrammed iCLMs and endogenous CMs. Overall, sarcomere organization was markedly lower in iCLMs than in CMs, indicative of less mature cardiomyocytes produced by GHMT reprogramming (**Figure 3.6A**). This result is consistent with our previous manual qualitative analysis, in which only ~20% of α-actinin^{+} cells were classifying as having visually distinguishable sarcomeres(Nam et al., 2014a). Though on average sarcomere organization was much lower in iCLMs, the presence of some iCLMs with highly organized sarcomeres indicates that this reprogramming method can produce cells on par with neonatal cardiomyocytes (**Figure 3.6D**). iCLMs with highly organized sarcomeres were found in α-actinin^{+} cells, α-actinin^{+}/Hcn4-GFP^{+} cells, and α-actinin^{+}/Hcn4-GFP^{+}/Nppa^{+} cells (**Supplementary Figure S3.3**). Sarcomere length was only accurately calculated for cells with sufficient sarcomere organization (sarcomere length >0.1). Cells with very low sarcomere organization scores would yield sarcomere length measurements well outside the normal range, possibly indicating other intensity-based features of the cells (**Supplementary Figure S3.4**). CMs with sufficient sarcomere organization for analysis had sarcomere lengths of about 2.2 μm, while iCLMs were typically lower, most of which fell in the 1.8–2.0 μm range (**Figure 3.6B**). Smaller sarcomere lengths are typically found in immature cardiomyocytes, again pointing towards a less mature phenotype(Yang et al., 2014b). Several α-actinin+/Hcn4-GFP+ and α-actinin+/Nppa+ iCLMs had very low sarcomere lengths of about 1 μm, which were manually confirmed in ImageJ (**Supplementary**

Figure S3.3). Surprisingly, these cells had very clear sarcomere structure, suggesting that this result may not be entirely due to an immature phenotype. Cell–sarcomere misalignment was also measured as the difference in angle between the orientation of the sarcomeres and the orientation of the major axis of the cell (**Figure 3.6C**). A smaller difference in angle, in which the sarcomeres are oriented in the direction of the major axis, is observed in mature cardiomyocytes(Bray et al., 2008). The cell–sarcomere misalignment was similar between CMs and iCLMs, however both cell types exhibited substantial variability. It should be noted that cells were not excluded based on cellular elongation. The orientation of the major axis in a less elongated cell is less meaningful, which may artificially result in a high misalignment score.

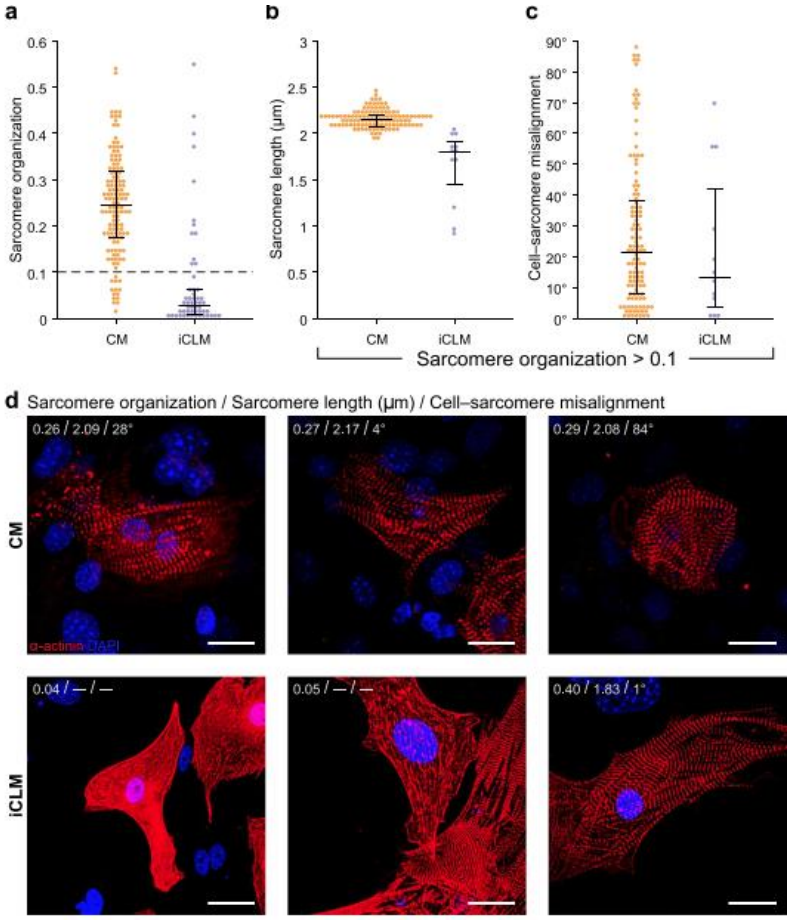


Figure 3.6. Sarcomere analysis of reprogrammed cardiomyocytes suggests an immature phenotype*. **A)** Sarcomere organization as calculated by Haralick correlation*

between neonatal mouse cardiomyocytes (CM) and fibroblast-reprogrammed induced cardiac-like myocytes (iCLM). Black bars represent 25th, 50th, and 75th percentile. Only cells with sarcomere organization >0.1 were analyzed for B) sarcomere length and C) cell–sarcomere misalignment. ***D)*** *Example images of sarcomere metrics in CMs and iCLMs. Sarcomere length and cell–sarcomere misalignment were not reported if sarcomere organization score was below the 0.1 threshold. Scale bar = 20 μm.*

The various morphological and cytoskeletal metrics were compared to identify potential relationships between different developmental operations (**Supplementary Figure S3.4**). In the CMs, there was a slight upward trend in cellular elongation with increasing sarcomere organization, which was not observed in the iCLMs. Above the sarcomere organization threshold, CMs exhibited no relationship between sarcomere organization and sarcomere length, possibly indicating that sarcomere length is no longer a major indicator of further cell maturity. Below the sarcomere organization threshold, a highly variable sarcomere length was observed in both endogenous and reprogrammed cardiomyocytes. This threshold effect suggests that sarcomeric proteins must be assembled in a highly coordinated fashion to tightly regulate the characteristic sarcomere length of a given cardiomyocyte. Cell–sarcomere misalignment was lower in the more elongated CMs, a relationship not seen in the iCLMs. This is possibly due to a coordinated effort in endogenous cardiomyocyte development to place new sarcomeres along the leading edge of the cell(Yang et al., 2014b).

DISCUSSION

These methods introduce a new framework for using multiple immunofluorescence channels to automatically segment cells and analyze both morphological and cytoskeletal

features of neonatal mouse and fibroblast-reprogrammed cardiomyocytes. The segmentation algorithm can be used without prior manual cell type classification, as it was with the reprogrammed cardiomyocytes. This allows for the identification of cells with any combination of cardiomyocyte markers, including the unexpected combinations we observed that are not seen in vivo.

Using our segmentation method, we found cells expressing cardiomyocyte markers similar to those of atrial cardiomyocytes (α-actinin$^+$/Nppa$^+$) and pacemaker cells (α-actinin$^+$/Hcn4-GFP$^+$). We have previously assessed GHMT-reprogrammed fibroblasts by patch-clamping(Nam et al., 2014a), and we found that induced atrial-like cells had action potentials similar to those of endogenous atrial cardiomyocytes. Similarly, induced pacemaker-like cells displayed action potentials resembling endogenous pacemaker cells. Although not assessed in this study, induced ventricular-like cells were also similar to endogenous ventricular cardiomyocytes. Morphologically, the α-actinin$^+$/Hcn4-GFP$^+$ cells resembled endogenous pacemaker cells in eccentricity and circularity, possibly suggesting a mature phenotype. To address the possibility that α-actinin$^+$/Nppa$^+$ iCLMs represented hypertrophic ventricular cardiomyocytes rather than atrial cardiomyocytes(Shubeita et al., 1990), we compared the α-actinin$^+$/Nppa$^+$ iCLMs to the α-actinin$^+$ cells, which may represent ventricular-like myocytes, and found no difference in cell size. Similarly, the α-actinin$^+$/Nppa$^+$ iCLMs are comparable in cell size to endogenous atrial cardiomyocytes. Although Hcn4 is dynamically expressed during heart development(Liang et al., 2013b), we have previously shown that these cells do not originate from an Nkx2.5$^+$ progenitor cell type or from actively dividing cells, suggesting

that these incompletely reprogramed cells arise from a committed lineage(Nam et al., 2014a).

Previous methods of assessing sarcomere structure are limited by the range of orientations and spatial patterns studied13 or require manual intervention(Aoki et al., 2000; Chopra et al., 2012; Pasqualini et al., 2015; Sheehy et al., 2014). Furthermore, our previous analysis of cardiomyocyte induction by GHMT relied on categorical classification rather than a continuous variable(Nam et al., 2014a). The high-throughput methods introduced here are fully automated and generalized to quantify sarcomere organization regardless of cell shape, cell orientation, and image magnification. We accomplish this by measuring Haralick correlation values at a wide range of orientations and pixel offsets. Non-regular geometries are tolerated because the Haralick correlation metric can be computed using only pixels belonging to cells. Fourier transforms, which are more commonly used, are limited to rectangular images.

To our knowledge, automated sarcomere texture analysis has not previously been applied to reprogrammed cardiomyocytes. Here, we apply our methods to both neonatal cardiomyocytes and fibroblast-reprogrammed cardiomyocytes. Using the described methods, we found that few α-actinin^{+} iCLMs had organized sarcomeres (~22%), compared to most endogenous CMs (~89%). In addition, iCLMs had shorter sarcomere lengths on average, which are typically seen in immature cardiomyocytes. Further, we found that cell elongation increases with sarcomere organization in the CMs but not the iCLMs. This suggests there may be a higher-level coordination between cell shape and sarcomere organization that has not been addressed in previous reprogramming studies. Indeed, it has been shown that, when cultured on micropatterned plates, sarcomeres

preferentially align with the major axis of elongated neonatal cardiomyocytes(Bray et al., 2008; Kuo et al., 2012).

Here we describe the utility of our algorithm for objectively evaluating cardiac reprogramming, but we envision that SOTA can be applied to additional research questions of considerable biological significance. For example, texture analysis could be similarly used to compare endogenous, reprogrammed, and iPS-derived cardiomyocytes, for which there remains concern about maturity and functionality(Yoshida and Yamanaka, 2017). Furthermore, SOTA could be used to track sarcomere formation in real-time in combination with appropriate fluorescent markers. Thus, we can foresee that such studies would allow more robust dissection of the molecular mechanisms that regulate sarcomere assembly(Sparrow and Schöck, 2009). From a more translational standpoint, real-time assessments of sarcomere formation could inform future screening efforts to optimize generation of functional cardiomyocytes. We also envision that our texture analysis approach could be applied to other muscle types, such as skeletal and smooth muscle, or even to other structurally complex cell types, such as neurons.

Several recent studies have described elegant and innovative whole-transcriptome approaches to characterize the potential functionality of specific cell types(Cao et al., 2016b; Zhang et al., 2015). However, we propose that gene expression signatures alone are unlikely to characterize the full range of intricately coordinated processes required for generating functional cardiomyocyte subtypes. For example, sarcomere gene expression does not guarantee efficient assembly and organization. It is likely that a combination of sarcomere protein stoichiometry, chaperone proteins, and specific post-translational modifications are required for proper sarcomere organization in addition to sarcomere

gene expression. Based on the potential importance of these non-transcriptional mechanisms, we believe that the sarcomere organization metrics described in our study will provide crucial information that remains uncaptured by current whole-transcriptome approaches. It is likely that combining the sarcomere analysis method with other approaches, such as electrophysiological and contractility measurements, will ultimately be required to function- ally optimize cellular engineering approaches for potential clinical translation.

METHODS

Isolation of mouse fibroblasts.

All experimental procedures involving animals were approved by the Institutional Animal Care and Use Committee at UT Southwestern Medical Center. All experiments and methods were performed in accordance with relevant guidelines and regulations. Mouse embryonic fibroblasts (MEFs) were isolated from E13.5 mouse embryos of Hcn4-GFP reporter or wild-type mice. The embryos were separated from the placenta and surrounding membranes. The head and the internal organs of the chest and abdominal cavities were removed from the embryos. The remaining tissues were minced and digested with 0.25% trypsin for 15 min at 37 °C to obtain single-cell suspensions. The isolated cells were cultured in fibroblast medium containing 10% FBS and 1% penicillin/streptomycin. These cells were trypsinized and replated the next day. Adult tail-tip fibroblasts (TTFs) were isolated using explant culture as described previously(Song et al., 2012a). Hearts from adult Hcn4-GFP mice were minced into small pieces which were

cultured in fibroblast medium. The medium was changed every 2−3 days. After ~10 days in culture, adult cardiac fibroblasts were harvested.

Isolation of CMs.

Neonatal mouse ventricular CMs were isolated using the Neomyts kit (Cellutron) as per manufacturer's protocol. Neonatal atrial and pacemaker CMs were isolated using methods modified from Sreejit et al.(Sreejit et al., 2008). P0−P1 hearts were dissected, washed in ice-cold PBS, and placed in Cold Balanced Solution (20 mmol/L HEPES 7.6, 130 mmol/L NaCl, 1 mmol/L NaH2PO4, 4 mmol/L glucose, and 3 mmol/L KCl). The right atrium was manually dissected and minced extensively in a minimal volume of 0.05% trypsin. Atrial tissue was incubated with 0.25% trypsin and agitated for 4 min in a 37 °C shaking water bath before allowing tissue to settle for 1 min without agitation. The first fraction was collected by removing the removing the supernatant to a fresh tube containing Culture Medium (DMEM, 20% FBS, 2 mmol/L L-glutamine, and 3 mmol/L sodium pyruvate). This cycle was repeated 3 times to collect a total of 4 fractions that were pooled, passed through a 100 μm filter, and combined with additional Culture Medium before plating on glass coverslips that had been previously coated with 2% gelatin for at least 10 min. After initial plating, the medium was changed after 72 h, and again every 48 h thereafter iCLM reprogramming. Generation of retroviral constructs of mouse Gata4, Hand2, Mef2c, and Tbx5 was performed as previously described(Song et al., 2012a). Retroviral constructs were transfected using Fugene 6 (Promega) into Platinum E cells (Cell Biolabs). 24 h after transfection, the viral medium (the media cultured with Platinum E cells) was collected and polybrene was added to viral medium

that was filtered through a 0.45 μm filter at a con- centration of 6 μg/μL. The mixture replaced the growth medium in the cell culture plate with mouse fibroblasts. Platinum E cells were replenished with the growth medium (DMEM with 10% FBS). 24 h later, mouse fibroblasts were re-infected with the second viral medium from Platinum E cell plate as described above for the first infection. Another 24 h later, viral medium on the plate with mouse fibroblasts was replaced with induction medium, composed of DMEM/199 (4:1), 10% FBS, 5% horse serum, 1% penicillin/streptomycin, 1% non-essential amino acids, 1% essential amino acids, 1% B-27, 1% insulin-selenium-transferrin, 1% vitamin mixture, and 1% sodium pyruvate (Invitrogen). 10% conditioned medium obtained from rat neonatal cardiomyocyte culture in DMEM/199 (4:1) with 5% FBS as described previously(Song et al., 2012a). Conditioned medium was filtered through a 0.22 μm filter. This medium was changed every 2–3 days until cells were harvested.

Immunocytochemistry.

Cells were fixed with 4% paraformaldehyde for 15 min and permeabilized with permeabilization buffer (0.05% Triton-X in PBS) for 5 min three times at room temperature. Cells were blocked with blocking buffer (Universal blocking buffer, BiogeneX) for 30 min and then incubated with primary antibodies against cTnT (Mouse monoclonal, Thermo Scientific, 1:400), α-actinin (Mouse monoclonal, Sigma, 1:400 dilution), GFP (Chicken IgY fraction, Invitrogen, 1:400 dilution), Nppa (Rabbit polyclonal, Abgent, 1:200 dilution), Myl2 (Rabbit polyclonal, Protein tech, 1:200 dilution), Myl7 (Rabbit polyclonal, Protein tech, 1:200 dilution), Myl7 (Mouse monoclonal, Synaptic system, 1:200 dilution), Hcn4 (Mouse monoclonal, NeuroMab, 1:200 dilution), PCM1

(Rabbit polyclonal, Sigma, 1:200 dilution) for 1 hr at room temperature or overnight at 4 °C (for Myl2, Myl7, Hcn4, and PCM1 antibodies). Following washing three times for 5 min with permeabilization buffer, cells were incubated with appropriate Alexa fluorogenic secondary antibodies (Invitrogen or Abcam) to detect the signal at room temperature for 1 hr. After another set of washing (5 min ×3 with permeabilization buffer), cells were mounted with Vectashield with DAPI and images were captured with Zeiss LSM 500 confocal microscope.

Cell segmentation algorithm.

Image analysis was performed on images from reprogrammed cells acquired in a previous study(Nam et al., 2014a) as well as images from new experiments using the same GHMT reprogramming methods and imaged as described above. Due to bleed-through of the Nppa and Hcn4-GFP channels into the DAPI channel, DAPI images were corrected by assigning pixels with >1.5 DAPI intensity to (NPPA or Hcn4-GFP) intensity ratio to zero. DAPI images were then smoothed using morphological closing followed by a Gaussian filter with a radius of 4 pixels. The image was then binarized using an Otsu threshold, which maximizes the variance between foreground and background pixel intensities. Small objects were removed from the image and nuclei were assigned an object number.

The Hcn4-GFP and α-actinin channels were similarly filtered with a Gaussian blur, and Otsu thresholding was done to produce binary images. The nuclei objects were overlaid with the Hcn4-GFP or α-actinin channel. Objects that were fully enclosed within the foreground of the image were classified as Hcn4+ or α-actinin+. If the objects were

fully within the foreground of the Hcn4-GFP/α-actinin intersection, they were categorized as Hcn4-GFP+/α-actinin+. Shortest distances were calculated between same-class nuclei, and nuclei within 25 pixels (~3.9 μm) of one another were treated as binucleates. Object numbers were subsequently reassigned.

To determine cell boundaries, α-actinin+ cells and Hcn4-GFP$^+$ cells were segmented first, in parallel. The Hcn4-GFP binary image was used to mask the α-actinin image and likewise the α-actinin binary image was used to mask the Hcn4-GFP image. Cells were segmented and these cellular regions were converted to masks to be applied to the combined Hcn4-GFP/α-actinin image. α-actinin$^+$/Hcn4-GFP$^+$ cells were then segmented in the masked image. In all cases, watershed segmentation was done on the gradient transformed image to determine cell boundaries(Vincent et al., 1991). Sobel horizontal and vertical edge-emphasizing filters were applied to the image and the magnitude of the two filtered images was taken. The marker-controlled watershed segmentation algorithm was used, treating nuclei as internal markers.

Because Nppa is perinuclear, it was not expected that Nppa signal would be high within the nucleus. Therefore, to classify cell objects for Nppa, a perinuclear ring was created for each nucleus. The perinuclear ring was defined as the area extending 8 pixels (1.25 μm) radially from the edge of the nucleus. Nppa intensity was measured in this region. If the 90th percentile intensity was greater than a manually determined threshold of 0.1 (relative intensity), cells were classified as Nppa$^+$.

Cell segmentation and sarcomere organization algorithms were written in MATLAB. A slightly modified cell segmentation algorithm was also developed in CellProfiler(Carpenter et al., 2006). This version uses manually determined mean

intensity thresholding in the nuclear area to classify cells instead of whole-image thresholding.

Cell morphology and sarcomere organization metrics.

MATLAB's image processing toolbox was used to compute cell size and shape characteristics. Cell area was computed in pixels and converted to μm(Prabhu and Frangogiannis, 2016). Cell elongation was calculated as the ratio of the major axis length to the minor axis length. Cell circularity is equal to $4\pi \times \text{Area} \times \text{Perimeter}^{-2}$, with a value of 1 indicating a perfect circle. Cell eccentricity specifies the eccentricity of the ellipse with the same second-moments as the cellular region. Eccentricity of the ellipse is calculated as the ratio of the distance between the foci and the major axis length. Cell eccentricity is 0 for a perfect circle.

Sarcomere organization is calculated using Haralick features, which are pixel intensity-based algorithms for quantifying image texture(Haralick et al., 1973; Haralick and Shapiro, 1992). First, a gray-level co-occurrence matrix is calculated for given orientation and pixel pair offset distances. The co-occurrence matrix p calculates the frequency at which pixels within a specified intensity range are matched by spatially separated pixels of the same intensity. The result is a g × g matrix, where g is the number of gray-levels (or intensity bins) that are to be considered. The default value was used, which for grayscale images is 8 (and 2 for binary images). From the co-occurrence matrix, 13 texture features can be measured. Haralick correlation is one of these features that

measures the likelihood of finding two pixels of similar intensity separated by a given distance. The correlation value is calculated by the Equation 1:

$$\text{Haralick correlation} = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j}$$

$$\mu_i = \sum_i i p_i(i) \qquad \mu_j = \sum_j j p_j(j)$$

$$\sigma_i = \sqrt{\sum_i (i-\mu_i)^2 p_i(i)} \qquad \sigma_j = \sqrt{\sum_j (j-\mu_j)^2 p_j(j)} \tag{1}$$

p(i,j) is the *i*th row and *j*th column of the co-occurrence matrix, and p_i and p_j are the marginal probabilities of the co-occurrence matrix.

Calculation of the co-occurrence matrix and corresponding Haralick correlation values was repeated at many orientation angles and spatial distances, resulting in an m×n matrix of Haralick correlation values, where m is the number of angles (values are symmetric about 180°) and *n* is the number of pixel offsets. We then used MATLAB's *interp1* function on each row to achieve sub-pixel resolution. Visualization of this data can be seen in **Figure 3.3**.

Other Haralick features that were considered in the comparison analysis include contrast (Equation 2), uniformity (Equation 3, also known as energy or angular second moment), and homogeneity (Equation 4), calculated from the co-occurrence matrix as follows:

$$\text{Haralick contrast} = \sum_{i,j} |i-j|^2 p(i,j) \tag{2}$$

$$\text{Haralick uniformity} = \sum_{i,j} p(i, j)^2 \tag{3}$$

$$\text{Haralick homogeneity} = \sum_{i,j} \frac{p(i, j)}{1 + |i - j|} \tag{4}$$

We also measured a variance value, which is the sum of the Haralick contrast values at an offset distance of 1 pixel. We expected the variance value to be close to zero for homogenous images and close to one for patterned images.

Gabor filters were applied at a variety of spatial frequencies and orientations using MATLAB's *imgaborfilt* function. The response magnitudes were normalized to cell area. Magnitudes were then plotted against wavelength for each orientation. Each profile was fitted to the sum of a quadratic function and a Gaussian function using MATLAB's *lsqnonlin* function to identify aperiodic and periodic components. A quadratic function was chosen instead of an exponential function because it appeared to fit better at longer wavelengths. The maximum amplitude of the Gaussian function was used as the Gabor filter score.

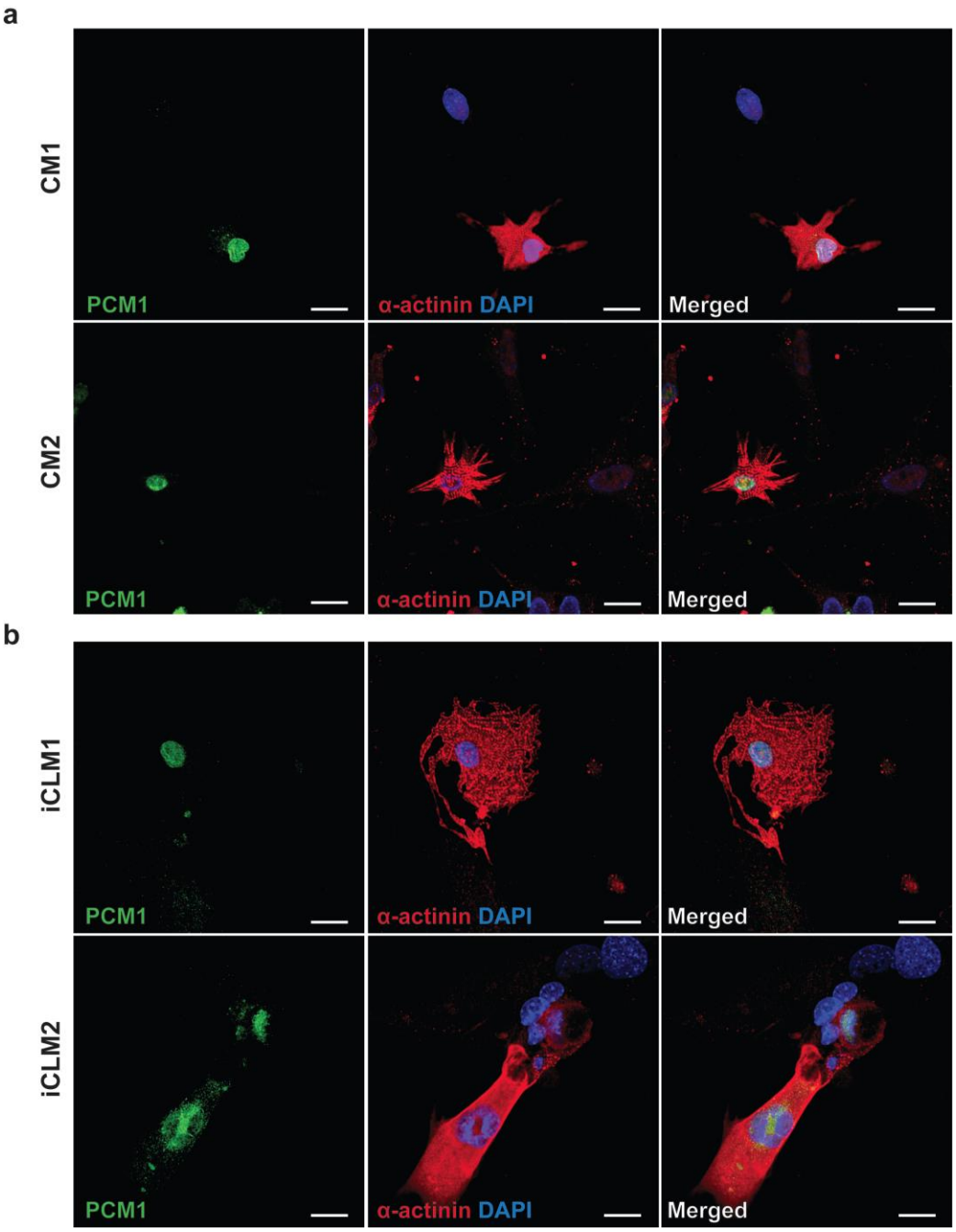
To generate Fourier transform scores, 2D Fast Fourier transforms were applied to the bounding box of the α-actinin channel, which produced a transformed image of the same dimensions. The subsequent analysis was done in a manner similar to that of a previously described method(Pasqualini et al., 2015; Sheehy et al., 2014). The transformed image was radially integrated along 360 dimensions to generate a frequency

profile. This profile was fit to the sum of two exponential functions and one Gaussian function using MATLAB's *lsqnonlin* function to identify aperiodic and periodic components. Sarcomere organization was calculated as the area under the Gaussian curve.

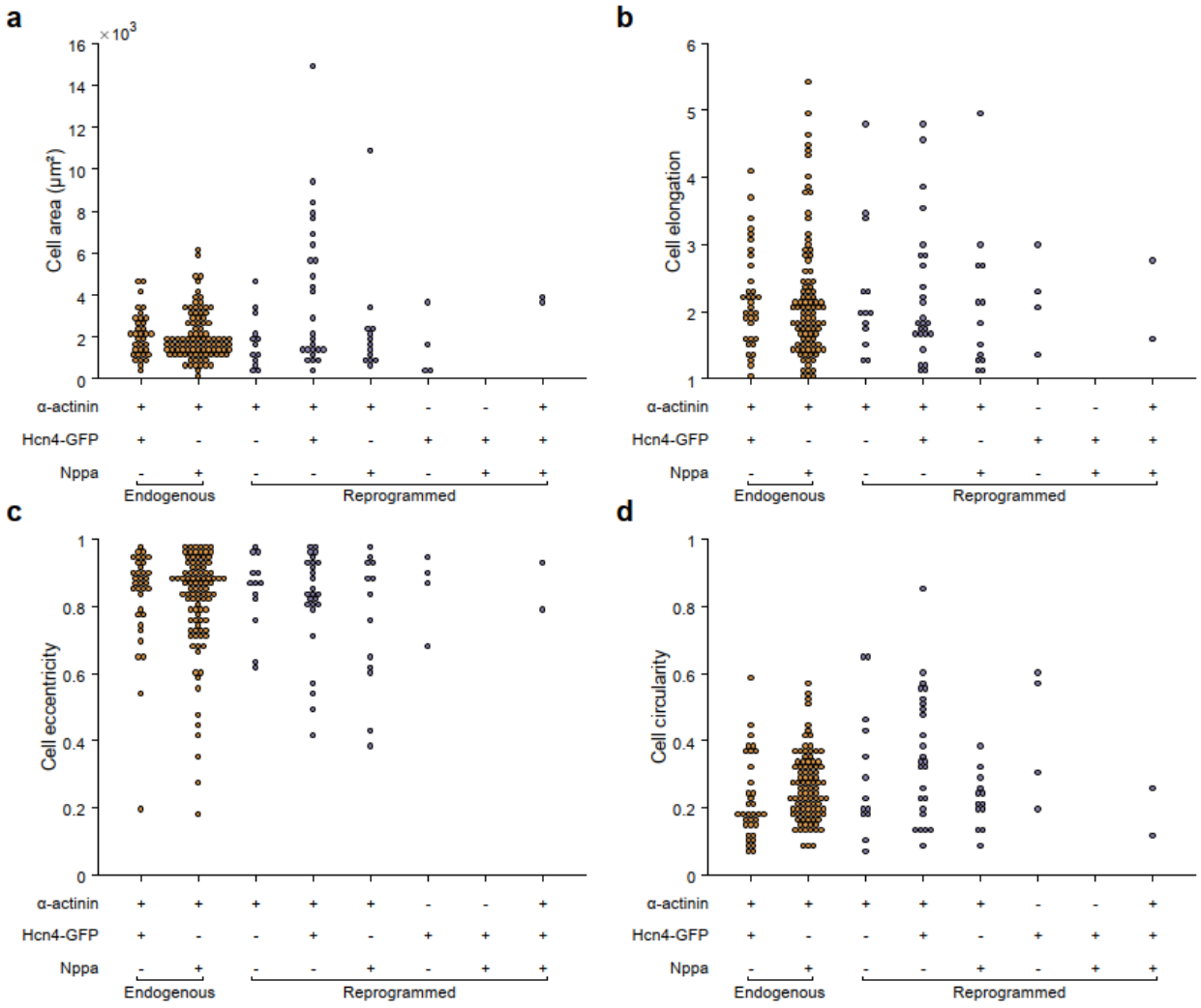
Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request. The MATLAB code and CellProfiler files used for the analysis are available at http://bme.virginia.edu/saucerman/.

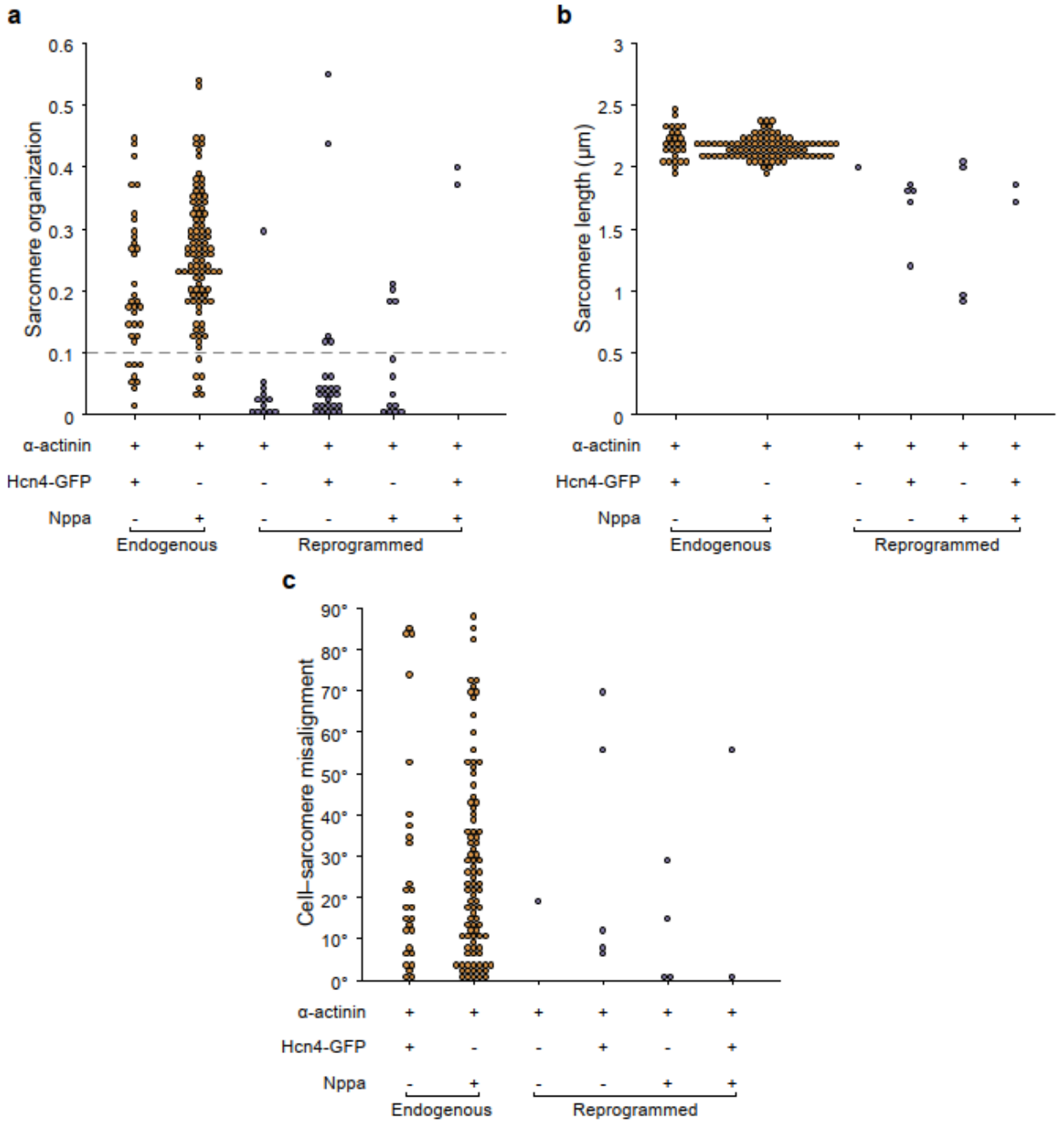
SUPPLMENTARY FIGURES



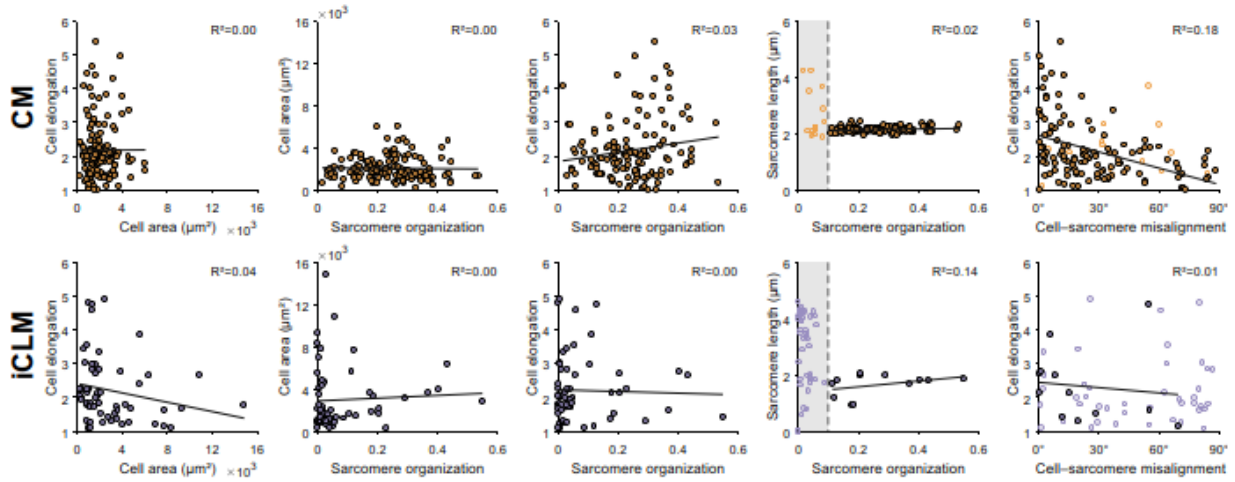
Supplementary Figure S3.1. Confirmation of cardiomyocyte induction and maturity by PCM1 staining. Representative. A) *endogenous CMs and* ***B)*** *reprogrammed iCLMs both express pericentriolar material 1 (PCM1) in the nuclear and perinuclear regions, suggesting mature cardiomyocyte induction(Bergmann et al., 2011; Zebrowski et al., 2015). In mitotic cells, PCM1 is found throughout the cytoplasm, associating with centrosomes during the cell cycle. PCM1 re-localizes to the nuclear envelope in post-mitotic cells, and this localization would be expected in terminally differentiated cardiomyocytes. Scale bar = 20 µm.*



Supplementary Figure S3.2. Cell morphology metrics by subtype. A) *Most of the variability in iCLM cell area is seen in α-actinin+/Hcn4-GFP+ cells. The variability in* ***B)*** *cell elongation,* ***C)*** *cell eccentricity, and* ***D)*** *cell circularity, is spread out equally among the different subtypes.*



Supplementary Figure S3.3. Sarcomere analysis metrics by subtype. A) *All iCLM subtypes were able to produce cells with sarcomere organization > 0.1.* ***B)*** *iCLMs with very low sarcomere lengths were either α-actinin+/Hcn4-GFP+ or α-actinin+/Nppa+.* ***C)*** *No clear pattern in cell−sarcomere misalignment among iCLM subtypes.*



Supplementary Figure S3.4. Scatter plots of cell morphology relationships. *CMs (top) and iCLMs (bottom) were measured for cell morphology and sarcomere organization. Open circles represent cells with below-threshold sarcomere organization.*

CHAPTER 3 ACKNOWLEDGMENTS

Support for this work was provided by the UVA Double Hoo Research Grant, the National Institutes of Health (HL05242 and HL135217), the Burroughs Wellcome Fund (#1009838), and the National Science Foundation (#1252854).

CHAPTER FOUR

AN ALTERNATIVE DELIVERY SYSTEM FOR CELL TRANSFORMATION

The contents of this chapter are reproduced from:

> Lam, H.K.*, **Fernandez-Perez, A., *** Schmidtke, D.W. and Munshi, N.V., 2018. Functional cargo delivery into mouse and human fibroblasts using a versatile microfluidic device. Biomedical microdevices, 20(3), p.52. *Primary Co-authors

Chapter Overview

The previous two chapters have focused in the discussion of iCM generation, quantification, and tool development of an automated quantification system. In this chapter I would like to shift focus and present our strategy to tackle another obstacle that we believe have a serious impact in reprogramming efficiencies. As shown in chapter two, the overall reprogramming efficiencies of well-organized iCM is about 1% of the initial infected population. Subtype diversity is then almost equally represented within that 1%. As you can appreciate, these numbers are troublesome for future clinical applications or more complex *in vitro* studies. However, as we discuss in chapter one, several groups have improved this efficiency by using additional transcription factors and/or small molecules. Nonetheless, this low efficiency brings several questions to mind 1) Is the low efficiency a fundamental aspect of fibroblast reprogramming? Or is it part of a more technical issue? Although I believe the answer lies between these two spectrums, I performed an experiment early in my career that highlights the importance of the latter. Given that I have to deliver four transcription factors (GHTM), we wondered what the probability of delivering all four factors to a given cell. For this, I generated four different

retroviral constructs expressing four different fluorescent proteins (TagBFP/GFP/TdTomato/E2-Crimson). I infected MEFs using a combination of these factors and quantified the probability of finding one, two, or more colors in each cell. In short, when infecting cells with one factor, the efficiencies are high (~90%). This almost halved when infecting with two different viruses, and it progressed to single digits when infecting with 3 or more constructs. This effect has been previously been reported in many studies (Superinfection), and it is probably a major contributor to our low efficiencies. As an attempt to bypass this issue, we decided to explore alternative forms of cargo delivery. Thanks to the close collaborations of UTSW and the UT Dallas bioengineering faculty, we were able to work with the Schmidtke laboratory in generating a microfluidic device capable of delivering multiple cargos of varying properties utilizing a cell deformation approach. Thus, in this chapter I will describe in detail the manufacturing and testing of Cyto-PDMS, a microfluidic device capable of delivering cargo of up to 70 kDa cargo. Kevin Lam and I prepared the manuscript, figures, and performed the experiments described in this chapter.

ABSTRACT

Efficient intracellular cargo delivery is a key hurdle for the translation of many emerging stem cell and cellular reprogramming therapies. Recently, a microfluidic-based device constructed from silicon was shown to transduce macromolecules into cells via shear-induced formation of plasma membrane pores. However, the scalability and widespread application of the current platform is limited since physical deformation-mediated delivery must be optimized for each therapeutic application. Therefore, we

sought to create a low-cost, versatile device that could facilitate rapid prototyping and application-specific optimization in most academic research labs. Here we describe the design and implementation of a microfluidic device constructed from Polydimethylsiloxane (PDMS) that we call Cyto-PDMS (Cytoplasmic PDMS-based Delivery and Modification System). Using a systematic Cyto-PDMS workflow, we demonstrate intracellular cargo delivery with minimal effects on cellular viability. We identify specific flow rates at which a wide range of cargo sizes (1-70 kDa) can be delivered to the cell interior. As a proof-of-principle for the biological utility of Cyto-PDMS, we show (i) F-actin labeling in live human fibroblasts and (ii) intracellular delivery of recombinant Cre protein with appropriate genomic recombination in recipient fibroblasts. Taken together, our results demonstrate that Cyto-PDMS can deliver small-molecules to the cytoplasm and biologically active cargo to the nucleus without major effects on viability. We anticipate that the cost and versatility of PDMS can be leveraged to optimize delivery to a broad array of possible cell types and thus expand the potential impact of cellular therapies.

INTRODUCTION

Recent advances in genome modification(Fellmann et al., 2017) and directed lineage conversion(Srivastava and DeWitt, 2016a) have transformed the future of disease-based therapies. To translate these amazing discoveries into human therapies, however, a major limitation remains safe and effective intracellular cargo delivery(Stewart et al., 2016). Currently, the most efficient vehicles for cargo delivery are viruses, and many virus-based gene therapies are actively being studied in human clinical

trials(Dunbar et al., 2018). Nevertheless, the long-term untoward consequences of viral delivery, including immune system activation, insertional mutagenesis, and systemic toxicity, remain to be characterized in detail(Dunbar et al., 2018). Furthermore, since viruses can only deliver nucleic acids, there will always remain a finite risk of genome insertion, even in episomal viruses such as adeno-associated virus(Kotterman and Schaffer, 2014). Therefore, many of these theoretical and observed adverse effects continue to motivate efforts to develop methods for non-viral intracellular cargo delivery.

The most prominent alternatives to virus-based cargo delivery are carrier-based and membrane disruption-based approaches(Stewart et al., 2016). Examples of carrier-based methods include ligand conjugates, cell penetrating peptides, nanoparticles, and various carrier materials(Stewart et al., 2016). Several studies have established the feasibility of these approaches, and a recent report demonstrated that lipid nanoparticles can package components of the CRISPR/Cas9 system to correct genetic hearing loss in mice(Gao et al., 2018). Carrier-based delivery systems seem well-suited for in vivo delivery into closed spaces (e.g. inner auditory canal), but in vitro studies have consistently demonstrated that endosomal escape remains a major barrier(Verdurmen et al., 2017). Thus, efficient methods for intracellular cargo delivery that provide access to the cytoplasm remain in high demand.

Membrane disruption-based cellular delivery can take many flavors, including both physical (e.g. heat, electroporation, etc.) and mechanical (e.g. shear, cavitation, hydrostatic, ballistic particles, etc.) methods(Takahashi and Yamanaka, 2006). Although many membrane disruption methods can cause protein denaturation or loss of intracellular contents, intracellular cargo delivery by cell squeezing has recently emerged

as a simple and efficient approach(Sharei et al., 2013). Cargo delivery by cell deformation is predicated upon the concept that the shear force generated by pushing cells through a constricted area creates transient plasma membrane pores that allow macromolecules to enter the cell by diffusion. Following intracellular delivery, membrane pores are rapidly sealed by endogenous repair mechanisms(Blazek et al., 2015), and the delivered cargo has direct access to the cytoplasm. Given the simplicity and efficiency of cargo uptake, cell squeezing is ideal for ex vivo delivery of a wide range of macromolecules, including proteins and impermeable small molecules.

Although the inherent versatility of cell squeezing has positioned this technology to substantially impact future development of cellular and genome modification therapies, current devices are restricted by two key features. First, since cell squeezing and pore formation rely upon generating precise amounts of shear force, devices must be optimized for individual cell types depending upon their size, shape, membrane composition, and mechanical properties. For example, constriction size is a key device variable that optimizes cargo uptake for individual cell types(Sharei et al., 2013). However, given the multitude of device operating parameters (constriction size and shape, inlet size and shape, outlet size and shape, device flow, etc.), we anticipate that additional features must be altered to optimize intracellular delivery in a cell type-specific manner. Second, most current cell squeezing devices are constructed from silicon, which is not readily amenable to rapid prototyping and optimization given the cost, expertise, and facilities required for their fabrication(McDonald and Whitesides, 2002). Furthermore, most academic biology labs do not have the capacity to generate and test many silicon-based cell squeezing devices for a desired application. For these reasons, we believe

that there is an unmet need for a low-cost, easily-fabricated cell squeezing device that is widely accessible to most academic research labs.

To facilitate rapid prototyping and optimization of cell squeezing for conversion of skin-derived fibroblasts into pacemaker cells(Nam et al., 2014b), we aimed to construct an alternative device from polydimethylsiloxane (PDMS). The advantages of using PDMS include low cost, minimal required expertise and resources, facile outsourcing for device construction, and rapid prototyping(McDonald and Whitesides, 2002). Furthermore, PDMS-based microfluidic devices can be readily sealed to glass slides to allow direct visualization and characterization of cells as they pass through the device. The use of PDMS would also provide the opportunity in the future to incorporate in-device cell culture after transcription factor delivery given the unique biocompatible properties of PDMS(Bhatia and Ingber, 2014; Luni et al., 2016). Although PDMS would thus seem to be the material of choice for constructing a cell squeezing platform, the high-pressure sensitivity of PDMS results in unpredictable flow behavior in microfluidic channels with high aspect ratios and has thus limited its use for applications that demand high flow rates(Gervais et al., 2006). Therefore, we sought to develop a novel PDMS microfluidic device that could withstand elevated flow rates to overcome this critical challenge.

In this study, we report the design and fabrication of Cyto-PDMS (Cytoplasmic PDMS-based Delivery and Modification System), a next-generation microfluidic device for intracellular cargo delivery via cell membrane perturbation. We describe a novel method for casting PDMS-based microfluidic devices and bonding them onto glass slides. Using this method, we show that Cyto-PDMS withstands high shear forces with minimal buckling in the constriction area. Furthermore, we show that Cyto-PDMS offers the key

benefit of directly visualizing cells as they pass through the constriction zone. Importantly, we demonstrate that Cyto-PDMS delivers a wide range of cargo sizes into the cytoplasm of human fibroblasts with minimal effects on cellular viability. Finally, we validate the biological utility of Cyto-PDMS by delivering a cell-impermeable actin-binding toxin to the cytoplasm and an enzymatically active Cre recombinase to the nucleus of live cells. Taken together, our studies establish Cyto-PDMS as a viable intracellular cargo delivery system that possesses unique and versatile design features for future scaling and rational device optimization.

MATERIALS AND METHODS

Wafer Fabrication

Cyto-PDMS was designed using a computer aided design (CAD) program (Autodesk, USA) with channel widths of 90 μm that truncate to a constriction width of 6 μm and a length of 30 μm (**Figure 4.1E**). **Figure 4.1A** shows an entire 45-channel Cyto-PDMS microfluidic device, and **Figure 4.1C** depicts the constriction zone along with the upstream and downstream areas. Silicon wafers (University Wafer, USA) were spin coated with KMPR 1005 (Microchem Corp, USA) at a speed of 1300 rpm and soft-baked for 5 min at 100°C(Shimp et al., 2016). Silicon wafers were then UV exposed at 335 mJ/cm^2 under a mask aligner (Karl Suss, Germany). Development of the wafers was completed using 20 mL of SU-8 developer (Microchem Corp, USA) with slight agitation for 2 min, rinsed under a stream of 2-propanol for 30 sec, and then completely dried using a stream of nitrogen. Wafers were then exposed to a silane treatment via degassing in a

chamber for 4 hours with Tridecafluor-1,1,2,2-Tetrahydroocytl Methyldichlorosilane (Gelest, USA) prior to casting.

PDMS Casting

PDMS (Polydimethylsiloxane) Sylgard-184 (Cat. No. 3097538-1004; Dow Corning, USA) was mixed at a ratio of 10:1 elastomer to curing agent, poured over the silicon wafers to a thickness of 1 cm, and degassed for 1 hour in a desiccator filled with silica beads. Metal rods (0.1" outer diameter, Amazon, USA) were then inserted into the inlet port of the device and oven-cured for 1 hour at 80°C. After the initial curing, the metal rods were gently removed, and 5 cm silastic laboratory tubing (Dow Corning, USA) was inserted into the cross-linked PDMS ¾ of the way down as shown at the bottom of **Figure 4.2C**. A second 1 cm layer of PDMS was cast around the tubing to seal it in place and oven cured overnight.

Chip Assembly and Testing

PDMS microfluidic devices were cut out of each wafer and then washed in the following solutions with sonication: 1 M HCl for 7 min; 100% Acetone for 5 min; 100% EtOH for 5 min; Millipore water for 5 min. PDMS devices were then oven dried at 80°C for 1 hour. A triple rinse of the microfluidic devices in Millipore water was performed following each washing step. PDMS and glass slides (VWR, USA) were exposed to air plasma (Harrick Plasma, USA) for 1 min at a pressure of 300 – 400 mTorr using the high radio frequency (RF) setting (18 W/cm^3) prior to bonding. Assembled chips were cured in

a desiccator for 7 days. Prior to device operation with cells, the chips were quality controlled leak tested by connecting the chips to a 1 cc plastic syringe (Covidien, USA) and perfusing PBS

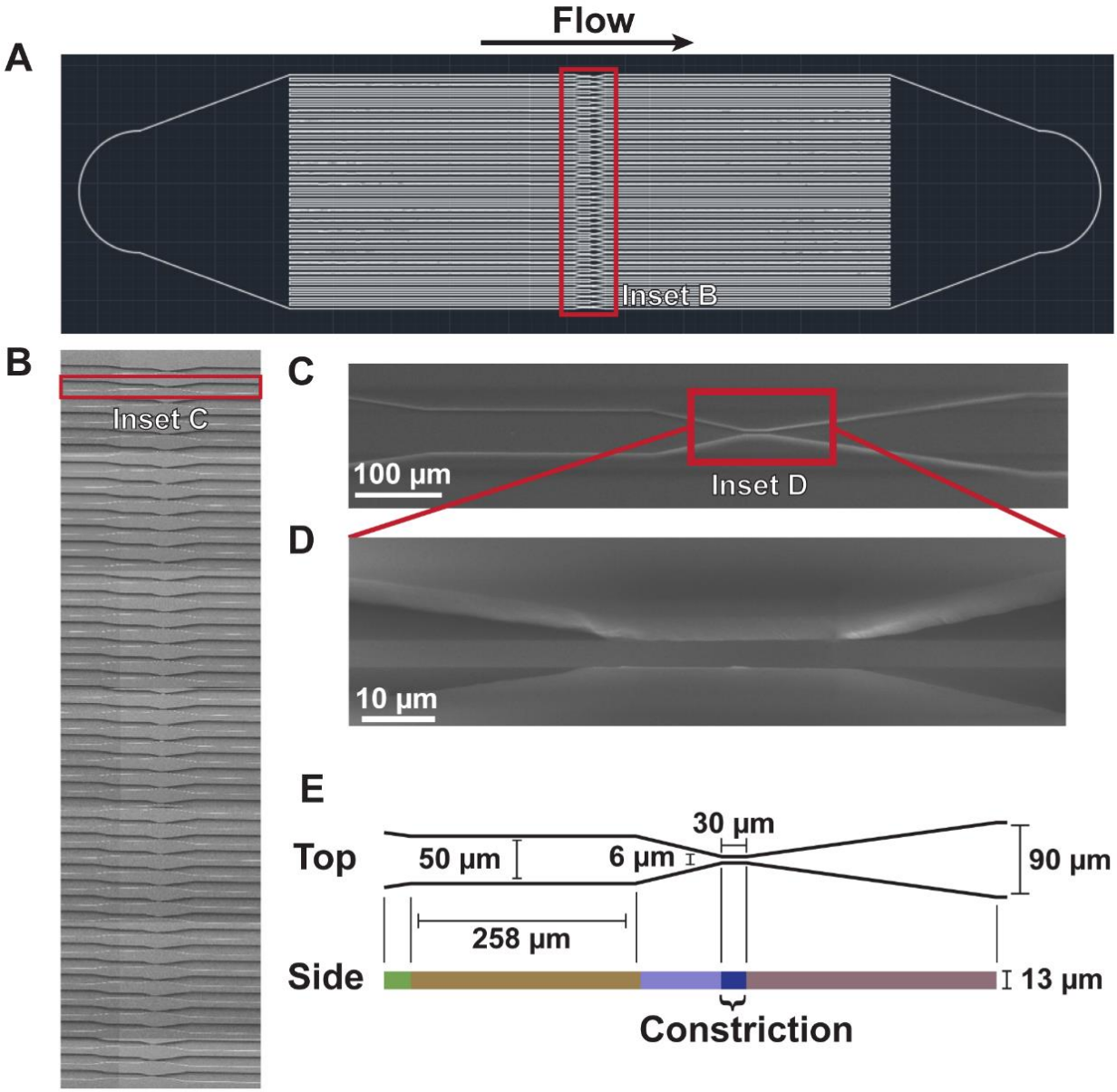


Figure 4.1. Cyto-PDMS device geometry and layout. A) *CAD design of Cyto-PDMS device showing 45-channels. The direction of flow is from left to right.* ***B)*** *SEM stitched image of device casted from PDMS showing all 45 parallel channels from inset B.* ***C)*** *ESEM imaging of single channel from inset C.* ***D)*** *Zoomed image shows the cargo-loading zone from inset D.* ***E)*** *Schematic of single constriction from device showing top down and side view with height of 13 µm. Device has channels that are 90 µm wide with truncation to 50 µm wide by 258 µm long leading to the cargo-loading zone that is 6 µm*

wide by 30 µm long. Side view is a representative image of the different sections of the device surrounding the cargo-loading area.

through the device with a syringe pump (KDS Scientific, USA) at a flow rate of 650 µL/min for 1 minute followed by perfusion of water (650 µL/min) for 1 minute. Residual water was removed, and the chips were stored in a desiccator until needed. Chips were considered functional if no leakage was observed through the entire device.

Channel Imaging

To image flow-induced deformation of the channel walls, a 3 µg/mL fluorescein isothiocyanate (FITC) (Thermo Fisher Scientific, USA) solution was infused through the device using a syringe pump (Harvard Apparatus, USA) at designated flow rates of 72.5, 150, 300, 450, 600, and 750 µL/min (n = 4 per flow rate) over a Zeiss Vert.A1 microscope using an A-Plan 10X objective (NA = 0.25) (Carl Zeiss Microscopy, Germany). Flow rates were allowed to stabilize for one minute prior to image capture using an Orca Flash 4.0 monochrome camera (Hamamatsu Photonics, Japan). One minute was chosen to simulate the minimum amount of time required for cargo loading experiments. Line intensity profiles upstream, at the constriction zone, and downstream of the constriction were quantified using the line profile function of ImagePro (Media Cybernetics, USA). Higher magnification imaging was also completed using an LD Plan-NEOFLAUR 40X objective (NA = 0.6) (Carl Zeiss Microscopy, Germany) and their respective line intensity profiles were also completed (**Figure 4.3**).

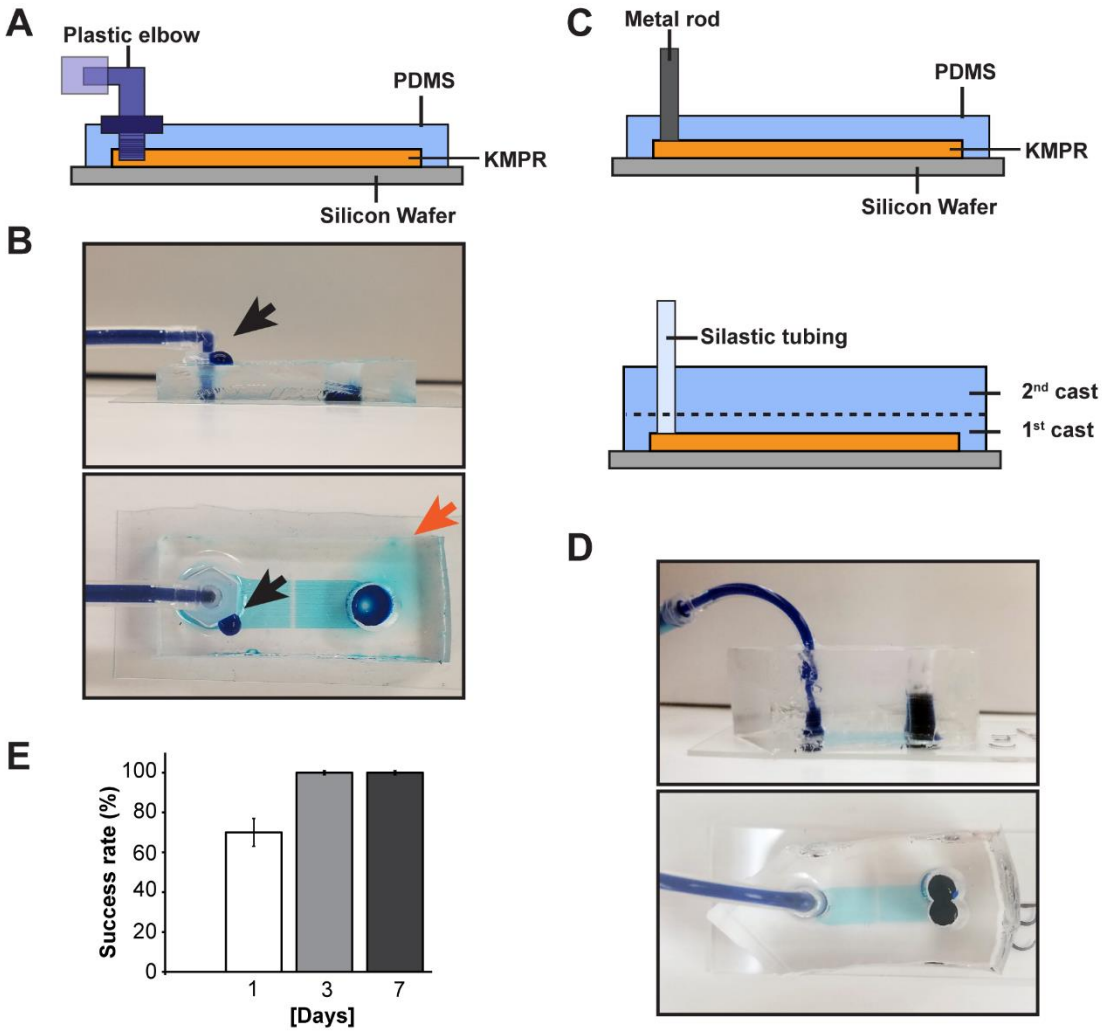


Figure 4.2. Schematic for Cyto-PDMS casting and bonding. ***A)*** *Side schematic of original Cyto-PDMS device. PDMS was poured over the photoresist and an elbow was inserted at the inlet.* ***B)*** *Top and side image of original concept for Cyto-PDMS device. Usage of an elbow in the inlet area lead to leakage (black arrowheads). Because the PDMS was also thinner, delamination of the PDMS to glass was observed (red arrowhead).* ***C)*** *Schematic for casting Cyto-PDMS device. Metal rod is placed in inlet port for first cast and cure. The rod is then removed and laboratory tubing is inserted in its place followed by a secondary cast and cure to seal the tubing in place.* ***D)*** *Top and side image of final Cyto-PDMS device. PDMS is much thicker, and elbow is replaced by tubing inserted into device and sealed in place.* ***E)*** *Bar graph showing the success rate of Cyto-PDMS devices cured for 1, 3, or 7 days (n = 9 devices per time point) and tested by infusion at 650 µL/min. Data in* ***E)*** *depicted as the mean ± s.d.*

Cell Lines and Culture Conditions

All media components were purchased from Sigma-Aldrich unless indicated otherwise. BJ-5ta human fibroblasts (CRL-4001; ATCC, USA) were used in cell-based experiments (except as noted) and cultured in a 4:1 mixture of Dulbecco's Modified Eagle Medium (DMEM) and Medium 199 supplemented with 10% (vol/vol) fetal bovine serum (FBS) and 1% (vol/vol) penicillin/streptomycin (Thermo Fisher, USA). For intranuclear Cre recombinase delivery **(Figure 4.8),** mouse embryonic fibroblasts (MEFs) derived from Gt(ROSA)26Sor$^{tm9(CAG\text{-}tdTomato)/Hze}$ mice (Jackson Laboratory, USA) were isolated as previously described(Fernandez-Perez and Munshi, 2017). MEFs were cultured in DMEM supplemented with 10% (vol/vol) FBS and 1% (vol/vol) penicillin/streptomycin. Two days before cargo delivery, 1x10^6 human fibroblasts or MEFs were seeded onto a 10 cm tissue culture dish. The day of the experiment, cells were washed twice with room temperature PBS (without calcium and magnesium), detached with 1.5mL of 0.25% trypsin/EDTA (Thermo Fisher Scientific, USA), and neutralized with DMEM/FBS media. Cells were counted, pelleted at 500 g for 5 minutes at 4°C, and resuspended in PBS at a final concentration of 1x10^5 per mL per sample.

Live-Cell Imaging

For brightfield imaging, BJ5-ta fibroblasts were detached, washed as previously described, and resuspended to a final concentration of 1.5x10^6 per mL in DMEM/FBS media. For fluorescent imaging, Fluo-4 (Thermo Fisher Scientific, USA) was diluted to 1 µM in HBSS without calcium and magnesium (Corning, USA), added to the BJ5-ta

fibroblasts, and incubated in a 37^0C waterbath for 30 minutes. Cells were then washed 3 times at 200 g for 10 min each in HBSS without calcium and magnesium. Cells were resuspended in HBSS with calcium and magnesium (Lonza, USA) to a final concentration of $1x10^5$ cells/mL and filtered through a 100 µm cell strainer (Thermo Fisher Scientific, USA) to create a single-cell suspension. Live-cell images were captured on a Zeiss Axio Vert.A1 microscope fitted with an EC Plan-NEOFLAUR 40X objective (NA= 0.75) (Zeiss Microscopy, Germany) at 1 frame/ms with a flow rate of 0.5 µL/min while fluorescent live-cell images were captured on a Zeiss Observer.Z1 with the same objective at 1 frame/ms and a flow rate of 5 µL/min. Images were processed using Zen Blue Software (Zeiss Microscopy, Germany).

Cargo Preparation

Dextran-fluorescent conjugates (Thermo Fisher Scientific, USA) were reconstituted in an appropriate amount of sterile water based on molecular weight: 100 mg/mL for 3kDa Dextrans (Cat No. D3306), 50 mg/mL for the 10kDa Dextrans (Cat No. D1820), and 25 mg/mL for the 40 and 70 kDa Dextrans (Cat No. D1845 and D1822, respectively). For F-actin labeling experiments, Alexa Fluor™ 488 Phalloidin (Thermo Fisher Scientific, USA) was reconstituted in methanol at a stock concentration of 6.6 µM. For Cre-mediated recombination experiments, 12 units of Cre recombinase (NEB, USA) were loaded directly per experimental condition.

Cargo Delivery

Cells and cargo were mixed together in 1mL of PBS without calcium and magnesium and perfused through the Cyto-PDMS chip at the desired flow rate. For Dextran conjugates, the samples were resuspended to a final cargo concentration of 0.1-0.3 mg/mL depending on size. To label the F-actin network, a final concentration of 0.15 µM Ph488 was used for each sample. After perfusion through the chips, cell samples were incubated for 5 minutes at room temperature to allow for membrane recovery. The cell samples were then transferred from the outlet to a new 1.5 mL Eppendorf tube, washed once with 1 mL of fresh media, and pelleted at 500 g for 5 minutes at 4°C. The cell pellet was resuspended in fresh culture media or FACS buffer (PBS, 5% FBS and 0.1% NaN_3) for further analysis. For imaging analysis experiments, cells were plated in 24-well plates containing coverslips that had been previously treated with 500 µl of SureCoat (Cellutron, USA).

Flow Cytometry

Cells were collected after cargo loading as described above. Each sample was resuspended in 350 µL of FACS buffer. Viable cells were analyzed by gating on cell size and granularity. To simultaneously assess small cargo delivery and cell viability, samples were loaded with Propidium Iodide (PI) and fluorescein diacetate (FDA) at final concentrations of 14.3 µM and 0.5 µM, respectively. For dextran uptake analysis, cells were also loaded with PI, and the efficiency of cargo uptake was calculated based on the PI^+Dextran-conjugate$^+$ population. Flow cytometry was performed on a FACSCalibur (BD

Biosciences, USA). Data was acquired using BD CellQuest Pro and analyzed with FlowJo software (Tree Star, USA).

Confocal Microscopy

To image the distribution of fluorescent cargo within cells, samples were fixed with 4% paraformaldehyde for 15 minutes at room temperature and permeabilized with 3 washes of 0.1% Triton X-100 for 5 minutes each at room temperature. Samples were mounted with Vectashield and DAPI (Vector labs, USA), and high-resolution images were captured with a Zeiss LSM700 confocal microscope using a Plan-Apochromat 20x/0.75 objective. To calculate the percentage of cargo positive cells on each coverslip, four regions of the sample were imaged per experimental condition. The total number of $DAPI^+$ and fluorescent $cargo^+$ cells was counted to calculate percent efficiency. Cre-mediated recombination efficiency was similarly calculated.

Statistical Analysis

Statistical differences were determined by Student's *t*-test using PRISM (GraphPad Software, USA). A P value less than 0.05 was considered to be statistically significant.

RESULTS

Cyto-PDMS Fabrication

We designed a microfluidic device (Cyto-PDMS) with high shear constricted regions for intracellular delivery of proteins and other macromolecules. The device was fabricated by standard soft lithography techniques in which a PDMS mold was produced from a photoresist master template. The PDMS stamp was irreversibly sealed to a glass coverslip by exposing both the PDMS and glass slide to an air plasma treatment prior to sealing. The design and working principle of the microfluidic device is illustrated in **Figure 4.1** and is similar to that proposed by Sharei and colleagues(Sharei et al., 2013). The device consists of an inlet channel, which then subdivides into 45 parallel constricted regions and subsequently recombines to a single outlet channel **(Figure 4.1A)**. Each parallel channel is 90 μm wide, truncates down to 50 μm wide by 258 μm long, leads up to the cargo loading zone at 6 μm wide by 30 μm long, and finally expands back to 90 μm wide **(Figure 4.1B-E)**.

When a solution of cells is perfused through the device, the cell must deform and stretch its membrane to traverse through the constricted region **(Figure 4.1C-D).** Cell shearing creates pores in the plasma membrane that allow molecules (i.e. proteins, growth factors, etc.) present in the perfusion media to passively diffuse into the cell. In contrast to similar devices in the literature(Sharei et al., 2013), we chose polydimethylsiloxane (PDMS) to fabricate the microfluidic chip due to its ease of fabrication, low cost, and transparency. In addition, PDMS can be easily peeled away from photoresist templates, thus allowing multiple devices to be fabricated from the same photoresist mold. For example, we have easily fabricated ~20 chips from a single mold.

Bonding Cyto-PDMS to Glass

Based on the flow-rates used previously for intracellular cargo delivery(Sharei et al., 2013), we anticipated that our devices would experience high pressures during cargo delivery. Therefore, a key aspect in the design of Cyto-PDMS was to develop a watertight, leak-proof device. The traditional approach of irreversibly sealing PDMS devices to glass is to expose both surfaces to an air or oxygen plasma immediately prior to bonding. Although this approach suffices for most microfluidic applications, the low shear modulus of PDMS makes the bonding between PDMS and glass substrates susceptible to failure under high-pressure flow(Inglis, 2010)**.** Initially, we used a plastic elbow as the interconnect for the device, exposed the glass and PDMS to an air plasma for 60 seconds, bonded the device, and allowed the device to cure for 1-2 hours at room temperature before testing (**Figure 4.2A**). However, when the perfusion rate increased from 72.5 to 600 µL/min, we observed delamination of the seal between the glass and the PDMS in the constricted region and leakage from the tubing interconnect at the inlet port to the microchannel **(Figure 4.2B).**

To address the problem of leakage from the tubing interconnect at the inlet of the microfluidic device, we eliminated the plastic elbow and developed a two-layered PDMS casting approach similar to that described previously for sealing silastic tubing to the inlet port **(Figure 4.2C)**(Wang et al., 2014b). The first PDMS layer (1 cm thick) was formed by pouring degassed PDMS over a photoresist template with a metal rod inserted into the inlet port. After removal of the metal rod and insertion of silastic tubing, a second layer of the same degassed PDMS mixture was poured over the entire device to completely cover the interconnect and to secure the tubing into place **(Figure 4.2D)**. The thickness of the

second layer was 1 cm, and the overall thickness of the two-layered PDMS device was ~2-2.5 cm. We found that PDMS devices with thicknesses less than 1 cm frequently failed when exposed to high-pressure flow.

To overcome PDMS delamination from the glass at high perfusion rates, we investigated the effect of longer PDMS-glass bonding times. A previous report demonstrated that the average burst pressure for PDMS (Sylgard 184) microfluidic devices bonded to glass steadily increased (140 kPa-544 kPa) as the room temperature bonding time increased from 4 to 27 hours.(Long-Fang Tsai, 2011) Thus, we hypothesized that longer PDMS-glass bonding times would produce devices that are more reliable. Two-layered microfluidic devices exposed to an air plasma for 60 seconds and bonded to glass for 1 day resulted in ~33% device failure, while increasing the bonding time to either 3 or 7 days led to ~100% device success **(Figure 4.2E).**

Device Characteristics

One limitation of PDMS is that it has a lower Young's modulus (1.32 to 2.97 MPa) compared to silicon (130 to 188 GPa) and thus is more susceptible to channel deformation under pressure driven flow(Hopcroft et al., 2010). Since changes in a channel's cross-sectional area would alter the shear profiles and the forces acting to deform cells, consistent performance of Cyto-PDMS requires minimal channel deformation during high-pressure flow. Therefore, we measured microchannel dimensions as a function of perfusion rate using a well-established fluorescence imaging assay(Hardy et al., 2009),(Kim et al., 2014). In this assay, an aqueous solution containing fluorescein

isothiocyanate (FITC) is perfused through the microchannels at varying flow rates, and FITC emission intensity is measured at various points along the device for each flow rate tested. Any changes observed in the FITC emission intensity with increasing flow rate correspond to channel deformation at a measured location (upstream, constriction, or downstream).

Figure 4.3A shows a representative fluorescent image of a microfluidic chip filled with FITC solution at a perfusion rate of 450 μL/min. Using this image, we measured fluorescence line intensity across five of the 45 constricted regions **(Figure 4.3B)** and found that the line intensity remained constant across the five channels at three different locations (i.e. upstream, constriction downstream). Similar measurements were collected at each of the three locations for flow rates ranging from 72.5 to 750 μL/min, and the average fluorescent intensity measurements across five different microchannels were plotted as a function of flow rate **(Figure 4.3C)**. Fluorescent measurements for the upstream location increased linearly with flow rate, suggesting that there was bulging of the microchannel ceiling in the upstream region due to elevated pressure. Similarly, fluorescent intensity measurements downstream of the constriction zone increased with flow rate, although the magnitude of increase was less than the upstream location. In contrast, fluorescent measurements within the constricted region changed minimally, suggesting that the cross-sectional area of the constriction zone does not expand with increasing flow rates. We repeated these studies at higher magnification to exclude the possibility that insensitive fluorescence detection accounts for the lack of change in fluorescence within the constriction zone, but we still found no significant change in

fluorescence within the constriction zone across a broad range of flow rates by imaging with a 40X objective **(Figure S1)**. Importantly, these observations suggest that the shear profile experienced by cells passing through the constriction zone do not change substantially even at high pressures, thus enabling our PDMS devices to deform cells consistently across a wide range of flow rates.

Aside from consistent performance across flow rates, we also wished to evaluate how Cyto-PDMS devices would respond to repeated cycles of low- and high-pressure flow. Previous work has shown that cells can be repetitively deformed at a given speed to improve cargo uptake(Sharei et al., 2013), and our data suggest that different cargo sizes can be optimally delivered at specific flow rates **(Figure 4.6C).** Thus, we envision that future chip designs may require cargo delivery by multiple cycles of varying flow rates. To directly test the performance of Cyto-PDMS following multiple rounds of high (750 μL/min) and low (72.5μL/min) flow rates, we conducted FITC perfusion experiments as described above. Even after two full cycles of high and low flow rates, we observed negligible variation in fluorescent intensity within the constriction zone (**Figure S2**). Therefore, we conclude that Cyto-PDMS maintains its mechanical properties even after *repeated cycles of low and high shear rates.*

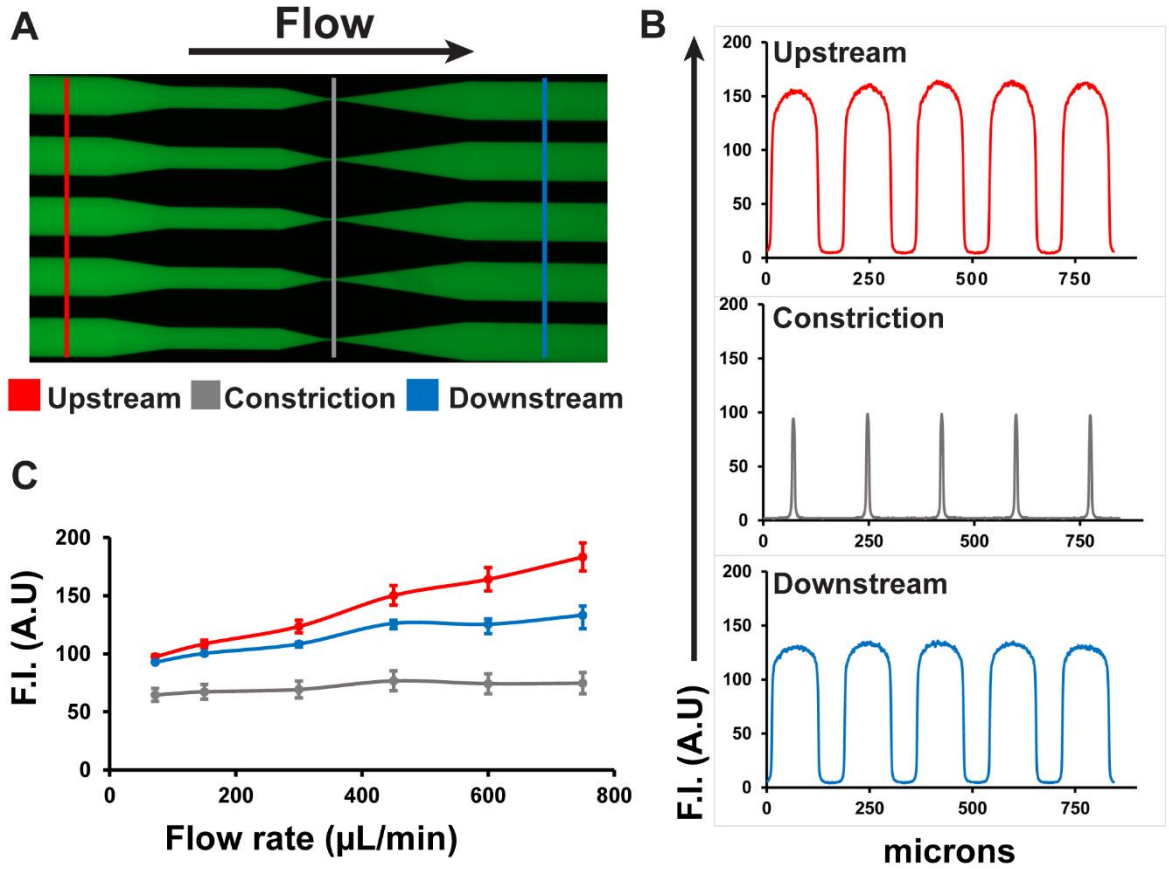


Figure 4.3 Minimal buckling in the Cyto-PDMS constriction area over a wide range of flow rates. A) *FITC was infused at 450 µL/ min into a Cyto-PDMS device. Intensity measurements across five constricted areas were taken at the upstream (red), constriction (gray), and downstream (blue) regions.* ***B)*** *Fluorescence Intensity (F.I.) measurements were taken across five channels from A) at the indicated regions. The vertical axis shows F.I. in arbitrary units (A.U.), while the horizontal axis represents the distance across the device in microns.* ***C)*** *Average F.I. of four independent experiments across all flow rates tested (72.5, 150, 300, 450, 600, and 750 µL/ min). Results in* ***C)*** *are depicted as the mean ± s.d.*

Live cell imaging with Cyto-PDMS

Since PDMS and glass are transparent, Cyto-PDMS devices provide an ideal platform for real-time imaging of cells passing through the constriction zone. Furthermore, the minimal auto-fluorescence of PDMS is compatible with fluorescent imaging(Cai et al., 2013). To directly observe how cells deform as they squeeze through the constricted

regions, we perfused solutions of fibroblasts ($1.5x10^6$ cells/ml) through a Cyto-PDMS device and imaged fibroblast movement through the channels by brightfield microscopy. As a fibroblast squeezed through the narrow constriction, its morphology changed from spherical upstream of the constriction **(Figure 4.4A**) to ellipsoid in the constriction **(Figure 4.4B-C)** and back to spherical **(Figure 4.4D)** downstream of the constriction **(Movie S4.1).** To examine the transit of cells through the constricted region by fluorescence microscopy, fibroblasts were labeled with the calcium-sensitive dye Fluo-4 prior to perfusion through the device **(Figure 4.4E-H,** and **Movie S4.2)**. Once again, we observed an ellipsoid morphology as the fibroblasts traversed the constricted region **(Figure 4.4F-G)**. Taken together, these results demonstrate that our PDMS-based devices provide the capability for real-time imaging of cellular deformation.

Intracellular Delivery by Cyto-PDMS

Having established that Cyto-PDMS can withstand high-pressure flow, we next wished to evaluate its ability to deliver cargo into cells. To accomplish this objective, we developed a perfusion paradigm adapted from previous work(Sharei et al., 2013). **Figure 4.5** schematizes our Cyto-PDMS experimental workflow. First, cultured fibroblasts are detached and washed prior to incubation with a solution containing the cargo of interest. Then, the cellular suspension is perfused through the inlet port of Cyto-PDMS, passed through the constriction zone, and incubated in the outlet port to allow for membrane recovery. Then, the loaded cells are washed and resuspended in an appropriate buffer.

Finally, the loaded fibroblasts are evaluated immediately by flow cytometry or plated, fixed, and immuno-stained on coverslips prior to analysis by confocal microscopy.

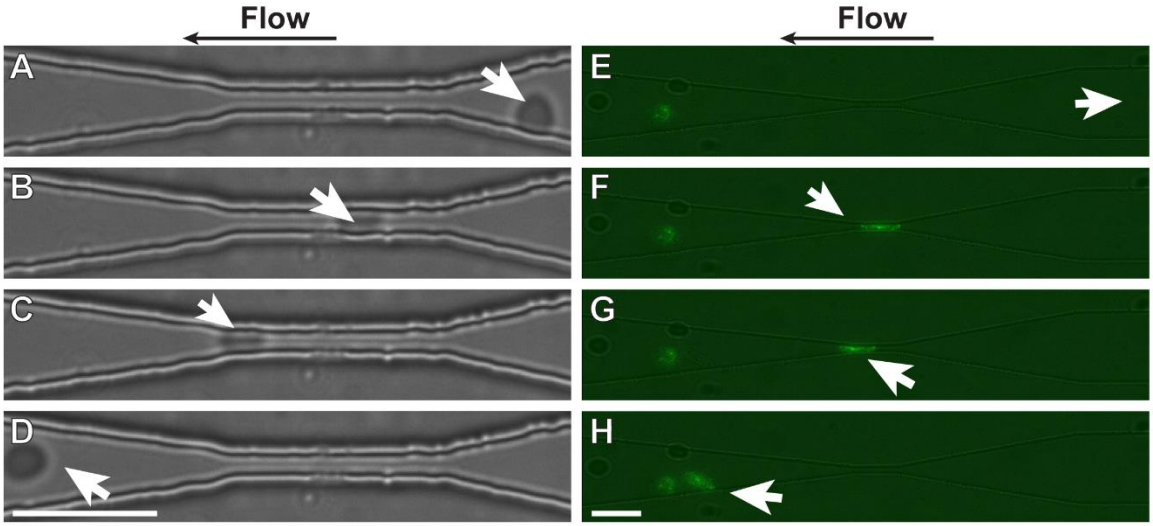


Figure 4.4 Live-cell imaging with Cyto-PDMS. A-D) *Live cells were imaged by time-lapse brightfield microscopy as they passed through a Cyto-PDMS device. Representative individual frames are shown for a single cell.* ***E-H)*** *Live cells were imaged by time-lapse fluorescence microscopy, and representative individual frames are shown. White arrowheads highlight a single cell as it transverses the cargo loading zone. Arrowhead in panel* ***E*** *points to the approaching cell. Scale bar = 20 µm.*

Maintaining high cell survival during intracellular cargo delivery is vital for future development of cell-based therapeutic applications. Therefore, we wished to simultaneously evaluate cargo loading and viability in loaded cells. To measure cargo uptake and viability, we adapted a previously described small-molecule loading assay(Fawcett et al., 1998). For this assay, the cell-permeable dye fluorescein diacetate (FDA) was used to confirm intact intracellular esterase activity. Only cells that are viable after cargo loading will hydrolyze FDA into fluorescein, which is detected by its green fluorescence **(Figure 4.6A).** In addition, fluorescein can only accumulate in cells with an intact plasma membrane. Thus, the presence of green fluorescence within a cell following intracellular FDA delivery demonstrates metabolic activity and an intact membrane, which

together indicate cellular viability. To model uptake of small cargo, we simultaneously loaded cells with FDA and propidium iodide (PI; MW = ~0.7kDa). PI is typically used to label DNA in dead cells because of its inability to cross the intact cell membrane of live cells. In the context of cargo delivery by Cyto-PDMS, however, PI can label DNA in live cells because transient membrane pores are created by cell shearing. Thus, we can measure the percentage of cells that simultaneously take up small cargo and maintain cellular viability by coupling FDA activity with PI inclusion in this assay.

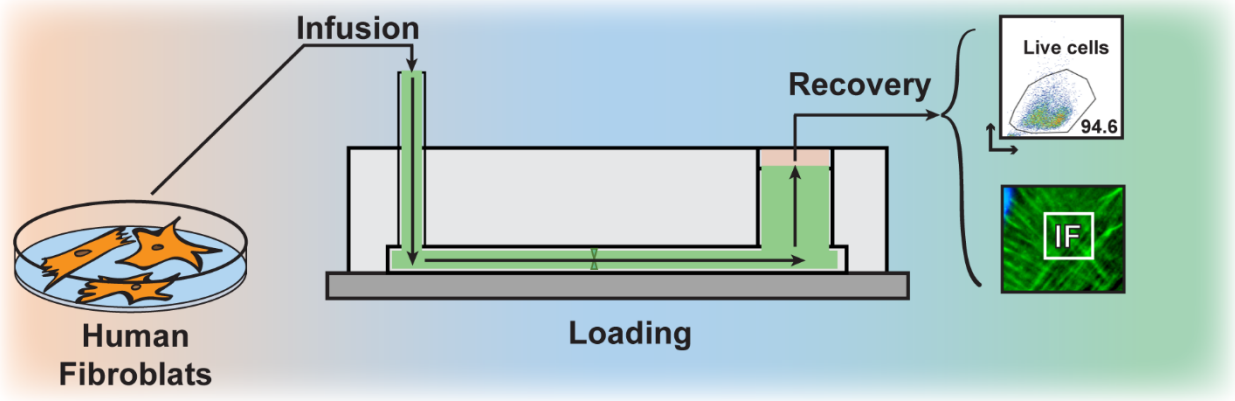


Figure 4.5 Workflow for intracellular cargo delivery by Cyto-PDMS. *Human fibroblasts are harvested, diluted in solution containing cargo, loaded into a syringe, and infused through the Cyto-PDMS. Cargo is loaded into the cells after pore formation, which occurs in theconstriction area. Loaded cells are incubated for 2 min at room temperature to allow for membrane recovery before further analysis by flow cytometry or immunofluorescence (IF).*

Initially, we sought to determine how perfusion flow rate influences small cargo uptake and cellular viability. Thus, we loaded cells with FDA and PI at flow rates from 72.5 to 750 µL/min **(Figure 4.6B**). Interestingly, we consistently found that >85% of loaded cells were viable and took up cargo at flow rates between 72.5 and 600 µL/min. Collectively, these data demonstrate that Cyto-PDMS can deliver small cargo and

maintain cellular viability across a broad range of perfusion rates. Since we observed a significant decline in cell viability at 750 µL/min **(Figure 4.6B**), we focused subsequent analyses on flow rates less than or equal to 600 µL/min.

One of the proposed applications for Cyto-PDMS is vector-free delivery of transcription factors for direct cellular reprogramming(Heins et al., 2002; Nam et al., 2013). Sharei et al. showed that it is possible to deliver purified Yamanaka factors (Pou5f1, Sox2, Klf4, and c-Myc) to generate induced pluripotent stem cells (iPSCs)(Takahashi and Yamanaka, 2006). However, iPSCs can renew indefinitely, while terminal cell types generated by direct reprogramming cannot proliferate, thus underscoring the importance of highly efficient transcription factor delivery to achieve clinically meaningful lineage conversion(Ieda et al., 2010b; Nam et al., 2014b; Vierbuchen et al., 2010b). In turn, efficient transcription factor delivery will critically depend upon cargo size and the operating features of the delivery system. For example, loading efficiency of CellSQZ is directly proportional to cargo size(Sharei et al., 2013). Thus, we sought to optimize delivery of a range of cargo sizes (i.e. 3-70 kDa) that are directly applicable to direct reprogramming.

We hypothesized that flow rate determines pore size and consequently affects cargo uptake in a size-dependent manner. For these experiments, we chose to use human foreskin fibroblasts (HFF), since they are capable of undergoing directed lineage

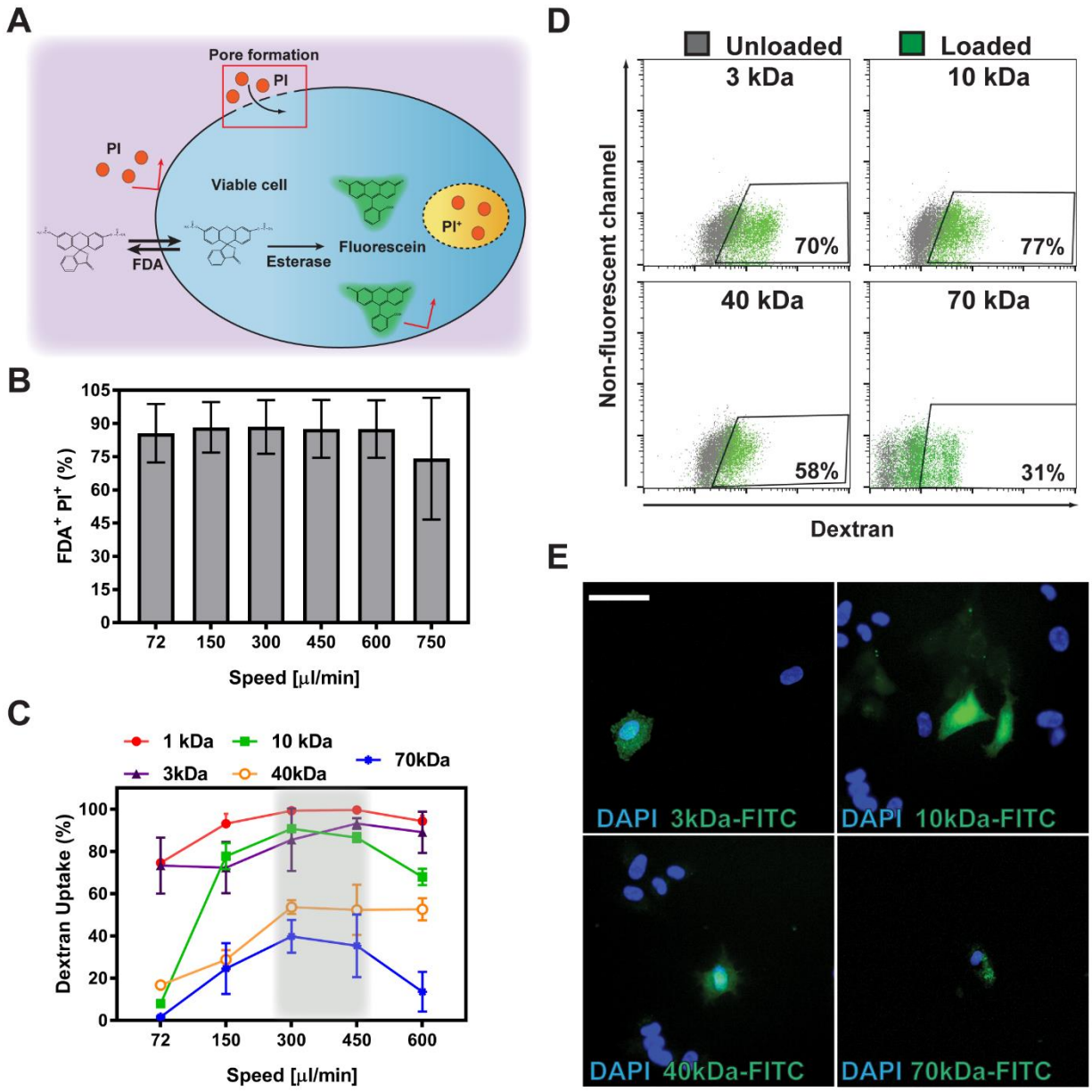


Figure 4.6. Flow rate influences cargo delivery efficiency. A) *Schematic for combined cargo uptake and cellular viability assay. PI is normally unable to enter live cells to label nuclear DNA. However, PI gains entry to the cell interior following Cyto-PDMS mediated pore formation. FDA is membrane permeant, but only viable cells can hydrolyze FDA into fluorescein. Once hydrolyzed, fluorescein is unable to leave cells with intact membranes.* ***B)*** *FACS analysis of FDA+PI+ BJ5-ta fibroblasts following Cyto-PDMS mediated delivery of FDA and PI at different flow rates (72.5 to 750 µL/min).* ***C)*** *Flow cytometry of cells with PI (1 kDa) and FITC-conjugated Dextrans of various sizes (3–70 kDa) loaded at different flow rates; shaded area (light gray) represents optimal flow rate range.* ***D)*** *Representative FACS data plots showing efficiency of cargo uptake at 450 µL/min. Unloaded control cells are shown in gray, while Dextran-loaded cells are shaded green.* ***E)*** *Representative confocal images of plated fibroblasts after loading with FITC-conjugated Dextrans at 450*

*μL/min. Scale bar = 50 μm. Results in **B** and **C** represent six independent experiments, and data is depicted as the mean ± s.d. PI, propidium iodide; FDA, fluorescein diacetate.*

conversion(Cao et al., 2016a; Nam et al., 2013). Based on the range of flow rates identified in **Figure 4.6B**, we loaded various-sized FITC-Dextran molecules (3, 10, 40, and 70 kDa) into cells by perfusing Cyto-PDMS devices at flow rates between 72.5 and 600 μL/min. The results of these studies, in addition to our experiments using PI (~1 kDa), are summarized in **Figure 4.6C**. Interestingly, we observed that small to intermediate-sized cargoes (1, 3, and 10 kDa) were effectively delivered (>75%) at flow rates between 150 and 450 μL/min. In contrast, larger-sized cargoes (40 and 70 kDa) never achieved the delivery efficiency of smaller cargoes. Nevertheless, 40 and 70 kDa Dextran delivery attained a maximum efficiency of >35% between 300 and 450 μL/min.

Figure 4.6D shows a representative set of flow cytometry plots from cells loaded with various-sized FITC-Dextrans (3, 10, 40, and 70 kDa) by perfusion through a Cyto-PDMS device at the optimal flow rate of 450 μL/min. Importantly, we confirmed that our flow cytometry data measured intracellular Dextran uptake, rather than non-specific adherence to the plasma membrane, through direct visualization by confocal microscopy of loaded cells **(Figure 4.6E**). Taken together, the results of these studies led us to conclude that flow rates between 300 and 450 μL/min achieve maximum delivery efficiency for a cargo size range of 1 to 70 kDa.

Given that the variation in delivery efficiency based on cargo size is not entirely dependent upon flow rate, we sought to identify additional characteristics that may explain the residual heterogeneity. One possibility is that the heterogeneity in cargo uptake correlates with an intrinsic property of the cells, such as cell size. To address this notion,

we evaluated the relationship between cargo uptake and cell size. We loaded human fibroblasts simultaneously with PI (~1 kDa) and FITC-Dextran (10 kDa). Then, we analysed the loaded cells based on cell complexity and size (**Figure S3A**). In parallel, we assessed the loaded cells for PI and Dextran uptake, and we identified three distinct populations: 1) $PI^{-}Dextran^{-}$, 2) $PI^{+}Dextran^{-}$, and 3) $PI^{+}Dextran^{+}$. Each of these individual populations was then back-gated onto the initial cell complexity/size analysis to identify unique features of each sub-population. This analysis revealed that $PI^{-}Dextran^{-}$ and $PI^{+}Dextran^{-}$ cells had similar mean cell sizes and size distributions **(Figure S4.3B).** In contrast, $PI^{+}Dextran^{+}$ cells had both a higher mean cell size and a narrower size distribution (**Figure S3B**). Thus, our flow cytometry analysis suggests that larger cells preferentially take up larger (10 kDa) cargo, while smaller cells tend to accumulate smaller (1 kDa) cargo, and the cells that take up the larger cargo demonstrate a more restricted size variance. Taken together, these observations suggest that cell size may dictate the efficiency of cargo uptake for a particular cargo size. However, this hypothesis remains to be formally tested. Nevertheless, optimization of additional factors, such as cargo concentration, pore formation, and pore closure, could further improve uptake of larger-sized cargoes in a cell type specific manner.

Cellular Uptake of Biologically Intact Cargo via Cyto-PDMS

Although our studies demonstrate intracellular Dextran delivery of various sizes by Cyto-PDMS, we sought to determine the feasibility of loading cells with biologically active cargo. Therefore, we tested the ability of Cyto-PDMS to deliver a cytoskeleton-binding toxin (Phalloidin) and a nuclear enzyme (Cre recombinase) into live cells.

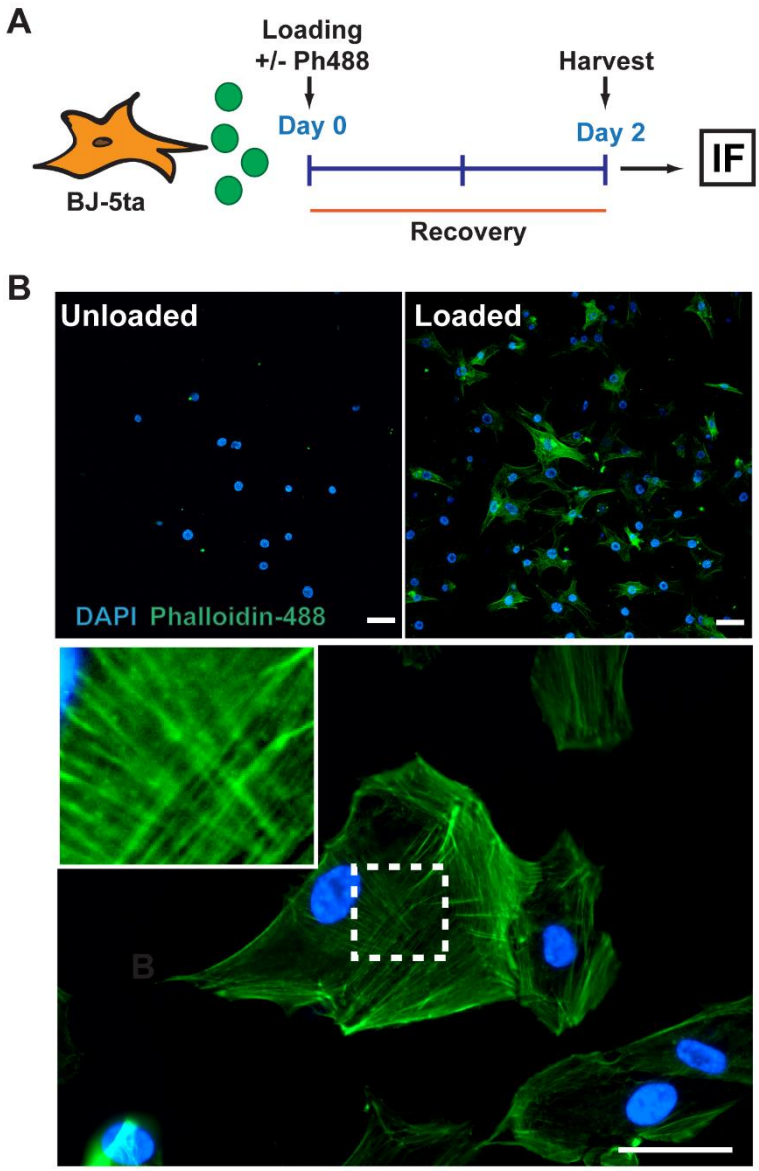


Figure 4.7 Cytoplasmic staining of the F-actin network via Cyto- PDMS cargo delivery. A) *Timeline for the loading of Ph488 into BJ5-ta human fibroblasts.* ***B)*** *Top: Representative confocal images of F-actin labeled cells with Ph488 at 450 µL/min 48 h post-cargo delivery. Scale bar 100 µm. Bottom: High magnification of a single BJ5-ta labeled cell. Scale bar = 50 µm. Inset: closer look at the F-actin striations. Ph488, phalloidin- Alexa488 conjugated (green); IF, Immunofluorescence.*

Intracellular mechano-transduction plays an important role in cellular reprogramming and differentiation(Connelly et al., 2010; Guo et al., 2014; Heisenberg and Bellaiche, 2013; McBeath et al., 2004). Aside from FRET-based measurements using

engineered biosensors(Meng et al., 2011; Meng et al., 2008), however, there has not been a consistent method to track cytoskeletal rearrangements in live cells. To address this gap, we wished to test the ability of Cyto-PDMS to label actin filaments in live human cells by loading BJ-5ta fibroblasts with the F-actin staining probe, Phalloidin-488 (Ph488) (**Figure 4.7A**). Phalloidin is a membrane-impermeable toxin enriched in *Amanita phalloides* mushrooms that selectively binds to F-actin(Baddeley et al., 2011; Liu et al., 2011). Typically, cells are fixed and permeabilized prior to labeling with Ph488. Using our Cyto-PDMS device, however, we were able to deliver Ph488 directly into the cytoplasm and label the F-actin network of living cells with high resolution **(Figure 4.7B).** Therefore, successful delivery of Ph488 with minimal toxicity demonstrates that the Cyto-PDMS cargo delivery system can be used to study actin network dynamics and potentially other intracellular structural properties in live cells.

To develop Cyto-PDMS as an eventual transcription factor delivery platform, it must be ultimately capable of transporting biologically intact cargo to the nucleus. As a proof-of-concept for intranuclear delivery, we sought to evaluate whether purified Cre recombinase protein could be loaded into fibroblasts by Cyto-PDMS and function within the nucleus. To address this question, we utilized primary mouse embryonic fibroblasts (MEFs) isolated from Cre reporter mice harboring a loxP-flanked STOP cassette that prevents transcription of tdTomato from the Gt(ROSA)26Sor locus(Madisen et al., 2010) **(Figure 4.8A).** If Cre recombinase is successfully delivered to the nucleus of these transgenic MEFs, then the Rosa26 locus undergoes recombination and tdTomato is expressed **(Figure 4.8A).**

Using this assay system, we wished to test the following: 1) whether Cyto-PDMS devices are able to deliver functional Cre protein into the cell interior and 2) whether intracellularly delivered Cre protein could reach the nucleus to recombine the ROSA26 locus. Thus, we used Cyto-PDMS to load transgenic MEFs with purified Cre protein and plated the loaded cells **(Figure 4.8B).** Successful Cre delivery and genomic recombination should result in robust tdTomato expression in a subset of plated cells. As expected, we observed tdTomato expression only in Cre-loaded cells **(Figure 4.8C)**. Quantification of tdTomato^{+} cells indicated that Cre uptake, nuclear transport, and genome recombination occurred in 14.5% of loaded cells **(Figure 4.8D)**. Altogether, this study clearly demonstrates that a ~40 kDa protein can be delivered intracellularly, migrate to the nucleus, and maintain sufficient bioactivity to recombine a genomic locus.

DISSCUSSION

In this study, we introduce a novel PDMS-based platform capable of withstanding high shear and pressure to deliver intracellular cargo. Although the transparency, versatility, and viscoelastic nature of PDMS have been studied previously(Anwar et al., 2011; Cai and Neyer, 2010; Cai et al., 2013; Hardy et al., 2009; Shiroma et al., 2016), we show that these properties can be leveraged to mediate intracellular cargo uptake. The flexibility of PDMS to conform to different structures with small feature sizes makes it an ideal candidate for this study. We have also observed that though PDMS has a much lower Young's modulus compared to silicon, the cargo-loading zone remains unaffected by high loading pressures **(Figure 4.3).** The PDMS platform can also withstand multiple cycles of high and low flow rates **(Figure S4.2)**, thus establishing its potential to

sequentially deliver cargo of varying sizes **(Figure 4.6).** These properties provide us with a versatile instrument that can be rapidly prototyped to optimize protein-mediated direct lineage conversion.

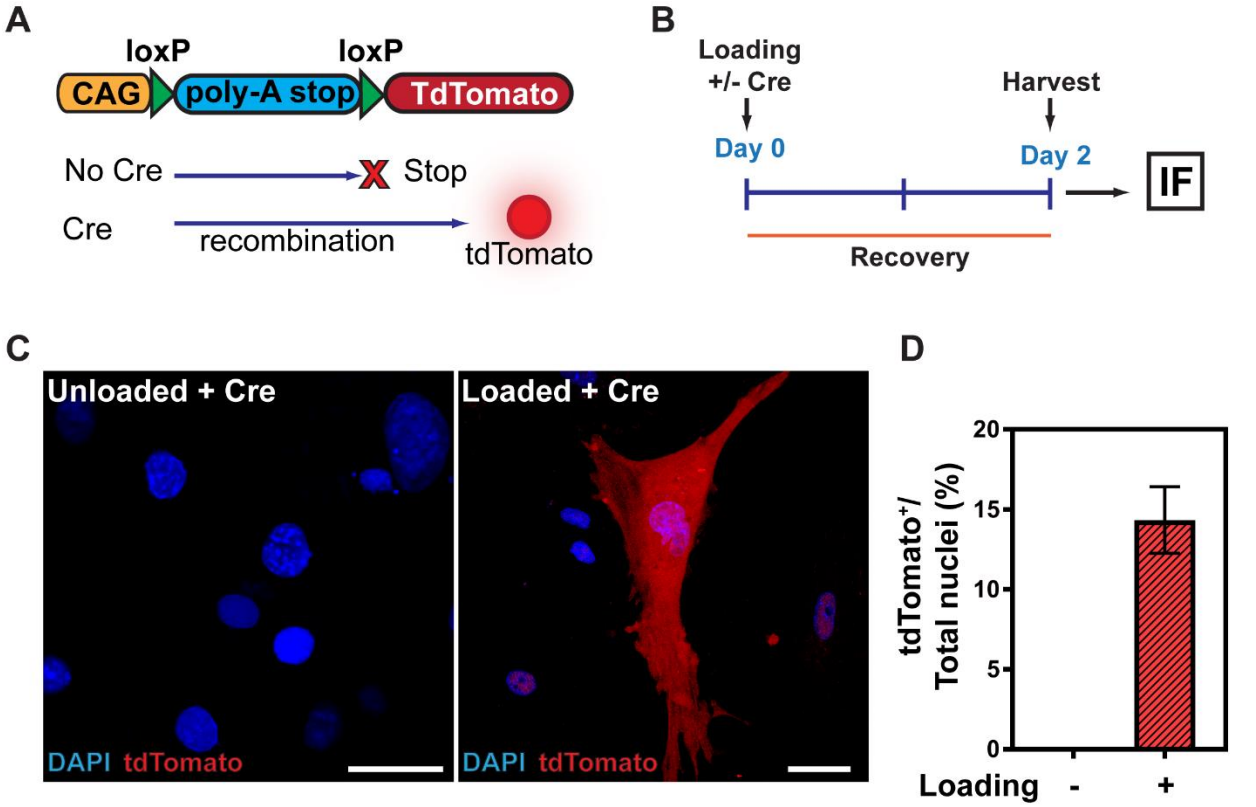


Figure 4.8 Nuclear delivery of biologically active cargo. ***A)*** *Illustration of the ROSA26-stopflox-tdTomato conditional knock-in allele. Without Cre, no recombination occurs. Upon addition of cre-recombinase, the poly-A stop is excised, and tdTomato expression is driven by the CAG promoter.* ***B)*** *Timeline for the loading of cre protein and harvesting 48h after cargo delivery.* ***C)*** *Representative confocal image of a recombined cell 48 h post- delivery of Cre (12 units) at 450 µL/min. Scale bar = 50 µm.* ***D)*** *Bar graph represents the count of tdTomato+ cells in six random imaging fields. Data in* ***D*** *depicted as the mean ± s.d. IF, immunofluorescence.*

Although Cyto-PDMS provides a promising avenue for rapid prototyping of future cellular delivery devices, the residual heterogeneity in cargo delivery requires further optimization. Among many potential explanations, we suggest that cargo-dependent cell size preferences **(Figure S4.3)** and inconsistent transit of cells through the constriction zone (data not sown) during Cyto-PDMS utilization may contribute to residual

heterogeneity in cargo delivery. Therefore, we must find ways to improve the consistency of cellular delivery, perhaps by making surface modifications to our device. Based on our preliminary data, we hypothesize that delivery of larger cargoes necessitates creation of larger membrane pores resulting from higher shear rates experienced at the cell surface. Thus, we plan to revise our device design in order to tune the shear force experienced by larger cells such that larger cargo uptake is favored. Alternatively, we could select cells based on size prior to exposing them to our Cyto-PDMS device. In this regard, PDMS-based devices that sort cells by size have been described(Yamada et al., 2007), and multiple PDMS devices can easily be connected in series.

As a proof-of-principle, Cyto-PDMS chips were shown to: (i) deliver a wide range of cargo sizes (1kDa-70kDa), (ii) maintain high cell viability (>85%; **Figure 4.6B**), and (iii) deliver active biological cargo **(Figure 4.8).** Furthermore, due to the transparent nature of both glass and PDMS (IOR of 1.5 and 1.4, respectively), a key advantage of the Cyto-PDMS platform is the ability to image live cells as they flow through the cargo loading zone(Cai et al., 2013; Whitesides and Tang, 2006) **(Figure 4.4).** Live imaging of cellular cargo uptake in the loading zone will enable future studies of shear effects on cellular physiology, cargo uptake kinetics, and membrane repair mechanisms. Given the ongoing need for vector-free intracellular protein delivery in potential therapeutic application of direct reprogramming and genome editing technologies, we envision that Cyto-PDMS will provide an alternative, cost-effective platform for mechanically loading cells.

SUPPLMENTARY FIGURES

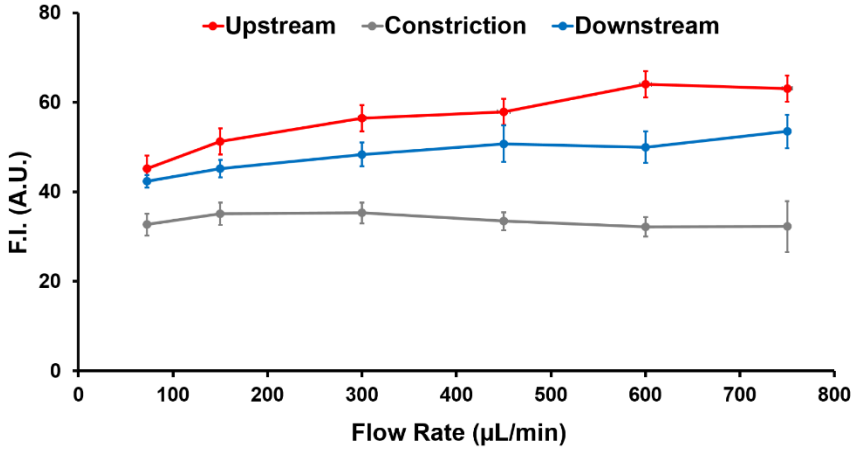


Figure S4.1. Fluorescence measurements and device deformation analysis are robust to magnification used for image capture. *Experiments were carried out as in Figure 4.3, except that images were captured at 40× magnification. Subsequent quantification of fluorescence in the upstream, downstream, and constriction areas demonstrated similar trends with varying flow rates. Consistent with the results shown in Figure 4.3, we observe no significant deformation within the constriction zone across a wide range of flow rates.*

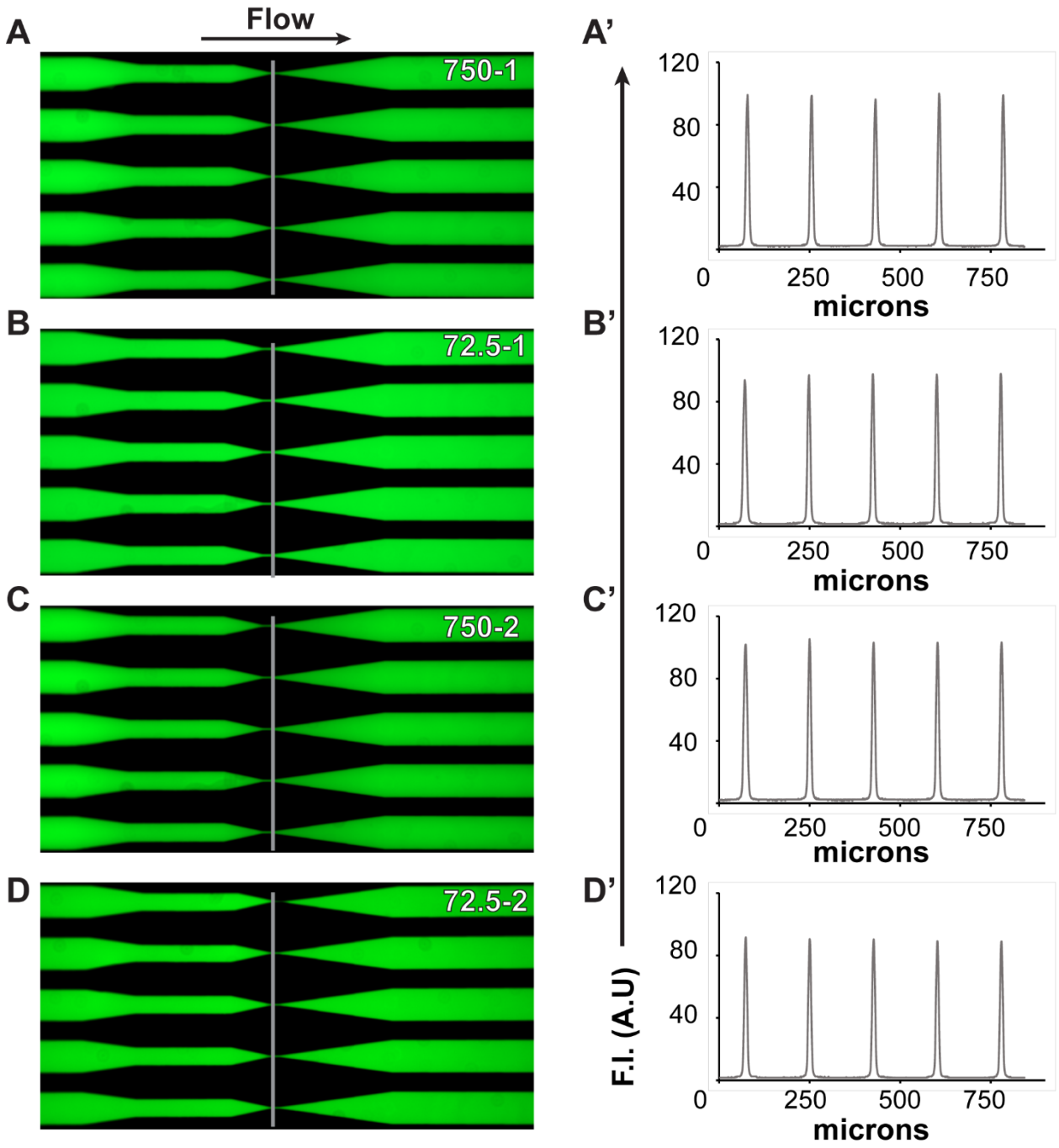


Figure S4.2. Cyto-PDMS withstands multiple cycles of changing flow rates. *Devices were tested with the highest and lowest flow rates to test the ability of PDMS to withstand changes while maintaining integrity. Flow rates were changed from 750 µL/min* ***A)*** *to 75.2 µL/min* ***B)*** *to 750 µL/min* ***C)*** *and ending with 72.5 µL/min* ***D)****. Fluorescence intensity measurements made at the cargo-loading zone did not show changes within the device due to the rapid change in flow rates* ***(A'-D')****. F.I, Fluorescent intensity; A.U., arbitrary units.*

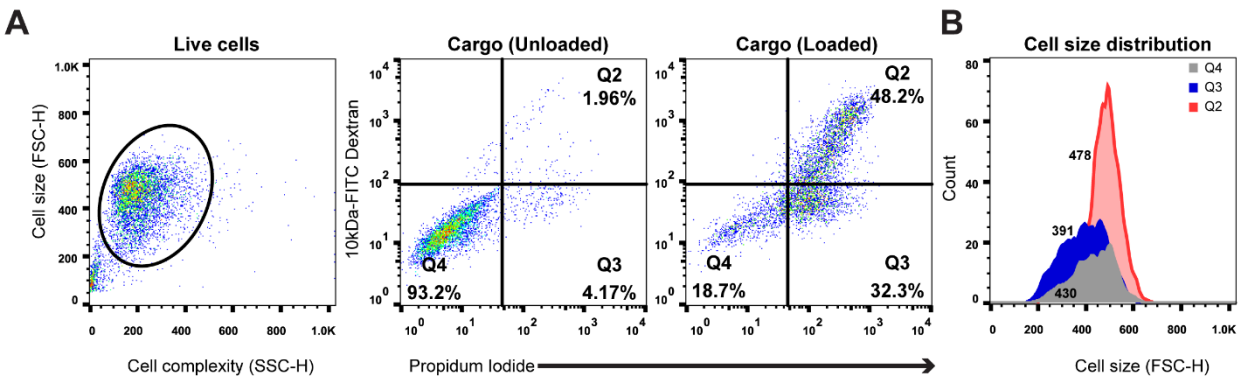


Figure S4.3 Retrospective analysis identifies cell size as a potential factor influencing cargo uptake efficiency. A) *BJ5-ta human fibroblasts were simultaneously loaded with propidium iodide (PI) and 10 kDa FITC-Dextran. After gating on live cells (left panel), mock-loaded and loaded cells were separated based on PI and Dextran uptake (right panels). Cells were then separated into 3 distinct sub-populations consisting of $PI^{-}Dextran^{-}$ cells in Quadrant 4 (Q4), $PI^{+}Dextran^{-}$ cells in Quadrant 3 (Q3), and $PI^{+}Dextran^{+}$ cells in Quadrant 2 (Q2).* ***B)*** *Cell size information for cells in each quadrant was then obtained and plotted as a histogram. While unloaded and PI-loaded cells are similarly sized, cells that also take up Dextran have a larger mean size and lower variance.*

CHAPTER 4 ACKNOWLDGMENTS

We thank members of the Munshi and Schmidtke labs for helpful suggestions and technical support. We gratefully acknowledge Nesreen Alsmadi for assistance with optical imaging of cells traversing through constricted channels. A.F.-P. was supported by the NSF (1000165336). D.W.S. was supported by funds from the Office of Vice President of Research at the University of Texas at Dallas. N.V.M was supported by funds from the NIH (HL094699, HL135217, and HL136604), the Burroughs Wellcome Fund (#1009838), and the March of Dimes (#5-FY13-203).

CHAPTER FIVE

HAND2 SELECTIVELY REORGANIZES CHROMATIN ACCESSIBILITY TO INDUCE PACEMAKER-LIKE TRANSCRIPTIONAL REPROGRAMMING

The contents of this chapter are reproduced from:

> **Fernandez-Perez, A.,** Sathe, A., A., Bhakta M., Legget, K., Xing, C., Munshi, V. N. Hand2 selectively reorganizes chromatin accessibility to induce pacemaker-like transcriptional reprograming. *Cell Reports.* 2019. In revision.

Chapter Overview

The contents of this chapter represent five years of reprogramming experiments. The initial observation that Hand2 enhances iCM formation and subtype diversity when added to GMT allowed me to dive-in and dissect a single transcription factor role in a broader paradigm. How is it that one transcription factor, that is not currently designated as a key regulator of the conduction system, allow for the formation of functional iPMs? In this chapter, I will describe a series of experiments that give insight to Hand2-GMT dependent chromatin remodeling of the mouse embryonic fibroblast toward functional iPMs. The computational analysis in this chapter was performed in collaboration with McDermott sequencing core.

ABSTRACT

Gata4, Mef2c, Hand2, and Tbx5 (GHMT) can reprogram transduced fibroblasts into induced pacemaker-like myocytes (iPMs), but the underlying mechanisms remain obscure. Here we explored the role of Hand2 in iPM formation by using a combination of transcriptome, genome, and biochemical assays. We found many shared transcriptional

signatures between iPMs and endogenous sinoatrial node (SAN), yet key regulatory networks remain missing. We demonstrate that Hand2 augments chromatin accessibility at loci involved in sarcomere organization, electrical coupling, and membrane depolarization. Focusing on an established cardiac Hand2 cistrome, we observe selective reorganization of chromatin accessibility to promote pacemaker-specific gene expression. Moreover, we identify a novel Hand2 cardiac subtype diversity (CSD) domain through biochemical analysis of the N-terminus. By integrating our RNA-seq and ATAC-seq datasets, we highlight desmosome organization as a hallmark feature of iPM formation. Collectively, our results illuminate Hand2-dependent mechanisms that may guide future efforts to rationally improve iPM formation.

INTRODUCTION

A growing body of evidence has established that cell identity is not completely fixed (Smith et al., 2016). Following the identification of transcription factor (TF) cocktails capable of generating specific cell types(Ieda et al., 2010; Takahashi and Yamanaka, 2006; Vierbuchen et al., 2010), transcriptomic and epigenetic studies have elucidated key mechanisms by which direct reprogramming is accomplished(Buganim et al., 2012; Liu et al., 2017; Treutlein et al., 2016). Based on these insights, reprogramming cocktails have been rationally improved in many instances to optimize formation of functional cell types(Buganim et al., 2012; Liu et al., 2017; Wapinski et al., 2017; Wapinski et al., 2013). Intriguingly, there are even select cases where detailed mechanistic insights into a particular reprogramming phenomenon has illuminated new developmental concepts(Mall et al., 2017; Pereira et al., 2016). While these lineage conversion

paradigms suggest plasticity in cell identity, it should be noted that many reprogramming systems used mouse embryonic fibroblasts (MEFs), which are inherently malleable.

Sinoatrial node (SAN) progenitors arise from the sinus horns to express Tbx5 and Tbx18 between E8.0 and E9.5(Mommersteeg et al., 2007). Interestingly, these progenitors are negative for Nkx2-5, which directly antagonizes the pacemaker gene program. During subsequent SAN specification, which takes place between E9.5 and E18.5, Shox2, Tbx3, and Isl1 participate in downstream regulatory networks. Shox2 antagonizes Nkx2.5 while activating Tbx3 and Isl1(Espinoza-Lewis et al., 2009). Tbx3 suppresses the gene expression program of neighboring atrial cardiomyocytes to demarcate the SAN boundary(Hoogaars et al., 2007). Finally, Isl1 directly activates the pacemaker gene program by cooperating with Shox2 and Tbx3(Liang et al., 2015). Our previous attempt to leverage this knowledge to reprogram pacemaker cells was unsuccessful(Nam et al., 2014). One possible explanation for this observation is that at least one pioneer TF is typically required to achieve somatic cell lineage conversion(Morris, 2016), yet these four factors function near the bottom of the pacemaker developmental hierarchy.

The combination of Gata4, Mef2c, and Tbx5 converts fibroblasts into functional induced cardiomyocyte-like myocytes (iCLMs)(Ieda et al., 2010). Addition of auxiliary factors (e.g. Hand2, Akt1, and Znf281) or manipulation of miRNAs, signaling pathways, culture conditions, or delivery strategies can further improve iCLM reprogramming (Abad et al., 2017; Ifkovits et al., 2014; Jayawardena et al., 2012; Miyamoto et al., 2018; Mohamed et al., 2017; Muraoka et al., 2014; Song et al., 2012; Yamakawa et al., 2015;

Zhao et al., 2015; Zhou et al., 2015; Zhou et al., 2017). We previously showed that GHMT can generate a subset of induced pacemaker-like myocytes (iPM) based on gene expression, flow cytometry, immunocytochemistry, morphological, and electrical characteristics(Nam et al., 2014). Moreover, we found that iPMs do not pass through an $Nkx2.5^{+}$ intermediate and exit the cell cycle rapidly, indicating that iPM formation is a direct reprogramming event(Nam et al., 2014).

Hand1/2 belong to the basic helix-loop-helix (bHLH) family of TFs and play essential roles during cardiac morphogenesis(George and Firulli, 2018; Wang and Baker, 2015). Hand2 is expressed in the embryonic right ventricle, while Hand1 is complementarily expressed in the left ventricle(George and Firulli, 2018), thus explaining the phenotypes of Hand1/2 knockout mice(Firulli et al., 1998; Srivastava et al., 1997). Hand2 also plays specific and important functions in the neural crest, epicardium, and endocardium. However, a role for Hand2 in SAN specification has not been documented to date. Thus, our prior observation that GHMT mediates iPM formation was surprising(Nam et al., 2014), and the underlying mechanisms by which Hand2 facilitates iPM reprogramming remain unclear.

Here we explored the role of Hand2 in iPM formation by using a combination of transcriptome, genome, and biochemical assays. We observed many shared transcriptional signatures between iPMs and endogenous SAN tissue of various stages, although specific key TF networks are missing. Focusing on the established Hand2 cardiac cistrome, we observe selective reorganization of chromatin accessibility that

promotes the activity of pacemaker-specific gene expression. We use fine-mapping of Hand2 protein domains to establish a novel cardiac subtype diversity (CSD) domain within the N-terminus. Finally, we identify desmosome organization as a key feature of iPM formation by integrating the RNA-seq and ATAC-seq datasets. Collectively, our results illuminate mechanisms by which Hand2 facilitates iPM formation and paves the way for future rational improvement of pacemaker formation.

RESULTS

Hand2 facilitates activation of cardiac conduction genes with GMT.

We first established an experimental system using Hcn4-GFP mice (Nam et al., 2014; Vedantham et al., 2015) to test mechanistic details of iPM reprogramming and a parallel pipeline for unbiased genome wide analysis of iPMs at day 14 (**Figure 5.1A**). To confirm that Hand2 is specifically required for iPM reprogramming, we conducted a series of experiments in which a single factor was systematically omitted from GHMT. MEFs were infected with specific TF combinations and assessed by flow cytometry at day 10 for Hcn4-GFP and cTnT (**Figure 5.1B**). We observed that GHMT generated 4.85% Hcn4-GFP^{+}/cTnT^{+} cells, which were not observed for any other combination.

Next, we used a more rigorous ICC assay to evaluate sarcomere formation and subtype diversification(Fernandez-Perez and Munshi, 2017; Nam et al., 2014). We previously confirmed the cell type identity of Hcn4-GFP+/sarcomere+ iPMs by direct intracellular recordings(Nam et al., 2014). Nevertheless, it should be noted that Hcn4 (and Hcn4-GFP) can label other cardiomyocytes of the conduction system aside from

pacemaker cells as well as certain non-myocytes, such as neurons, which do not express cardiac α-actinin or exhibit organized sarcomeres. Consistent with previous studies(Nam et al., 2014; Song et al., 2012), we found that Hand2 augments sarcomere formation and generation of all cardiac subtypes when added to GMT (**Figure 5.1C**). GHMT were all required for iPM formation, while Gata4 and Hand2 appear to play somewhat redundant roles during iAM and iVM formation, as GMT and HMT produced similar results in this reprogramming assay. Using an mCherry-Hand2 fusion construct, we conducted parallel studies of GHMT reprogramming that demonstrated that 80% of iCLMs and 100% of iPMs are Hand2-positive (**Figures S5.1A-B**). Altogether, these results show that Hand2 is required for iPM reprogramming in the context of GHMT.

To gain a more comprehensive understanding of the transcriptional changes that underlie iPM reprogramming, we performed RNA-seq analysis on transduced MEFs at day 14. We were particularly interested in the unique ability of Hand2 to generate iPMs with GMT, so we profiled GFP^+ cells after transduction of Hcn4-GFP MEFs with GMT or GHMT. In contrast to prior studies, we note that our transcriptome profiling occurs later in the reprogramming process, uses GFP-sorted cells, and directly compares GMT- and GHMT-transduced fibroblasts(Ieda et al., 2010; Song et al., 2012; Zhao et al., 2015). RNA-Seq was performed in duplicate with good overall concordance between samples (**Figure S5.1C**). Unbiased analysis of the RNA-seq data revealed four major clusters (1-4) of gene expression patterns (**Figure 5.1D**). Analysis of each cluster in aggregate allowed better visualization of quantitative changes in gene expression amongst samples (**Figure 5.1E**).

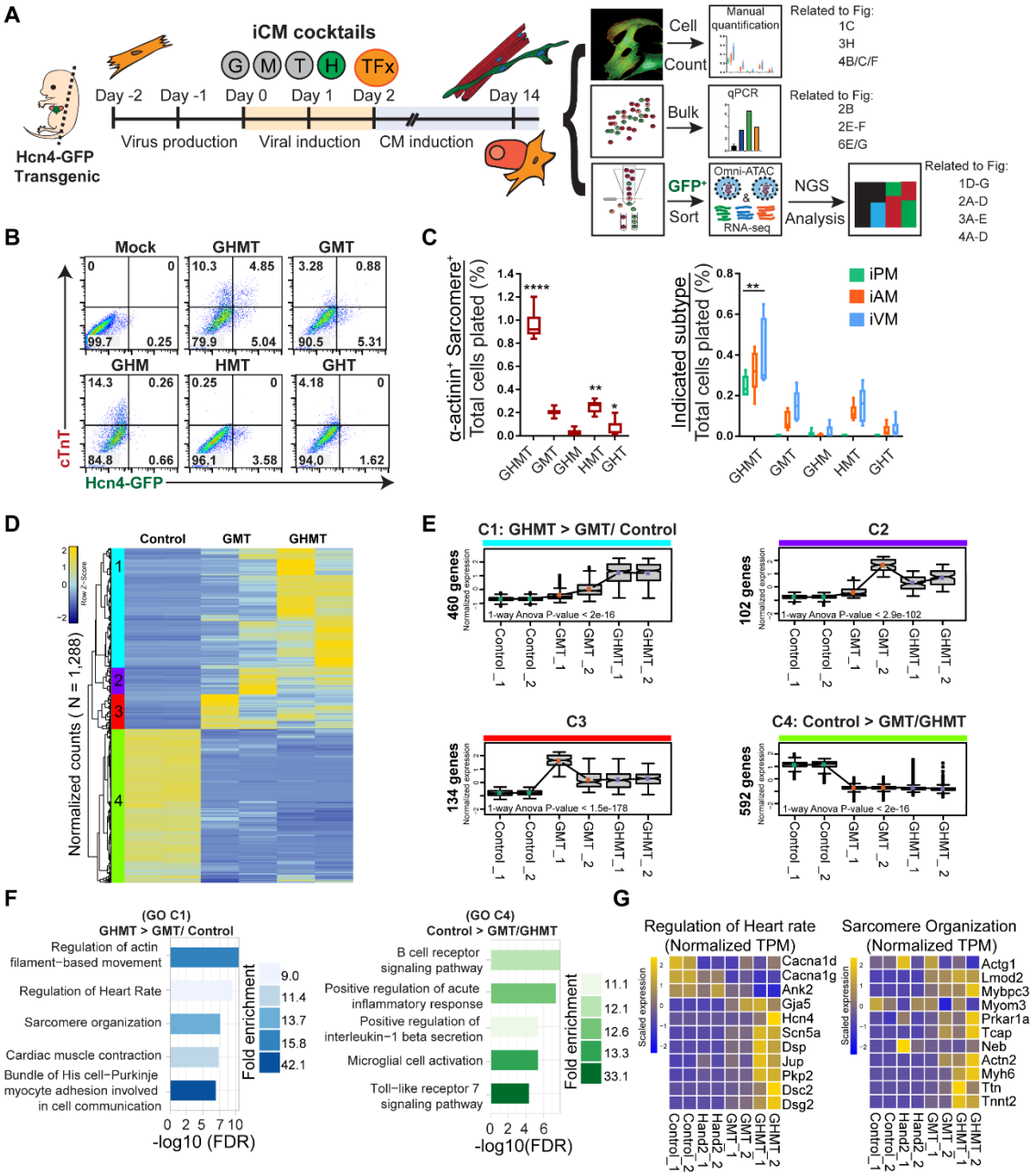


Figure 5.1. Hand2 facilitates transcriptional reprogramming of induced pacemaker (iPM) cells. *A) Experimental scheme for iPM reprogramming. B) MEFs were transduced with GHMT or each three-factor combination followed by flow cytometry analysis at day 10. C) Same experimental setup as in (B) but assayed by ICC and manual counting (mean±SD, n=6). D) RNA-seq analysis was performed on control, GMT-infected, and GHMT-infected MEFs. Unsupervised hierarchical clustering of these genes highlight four distinct clusters (1-4). E) Gene expression for each cluster in (D) is shown for individual samples. p-values for 1-way ANOVA testing across samples are shown within individual box plots. Statistical testing for pairwise comparisons within a given cluster was performed by Tukey's HSD test between groups. All comparisons were significant except for GHMT vs. GMT in Clusters 2 and 4. All numerical values for Tukey's HSD testing are*

provided in Table S2. F) Gene ontology (GO) analysis for clusters 1 and 4. G) Gene-level view of RNA-seq data comprising "sarcomere organization" and "regulation of heart rate" GO terms across samples.

Based on the coordinated changes in gene expression and the consistency of the data (**Figures 5.1E** and **S5.1D**), we focused on clusters 1 (C1: GHMT>GMT/Control) and 4 (C4: Control>GMT/GHMT) for gene ontology (GO) analysis. For GHMT>GMT/Control (C1), we observed enrichment for the expected categories "regulation of cardiac muscle contraction" and "sarcomere organization" (**Figure 5.1F**). However, we also identified the GO terms "regulation of heart rate" and "Bundle of His cell-Purkinje myocyte adhesion involved in cell communication," suggesting activation of conduction system-specific genes (**Figure 5.1F**). For Control>GMT/GHMT (C4), GMT and GHMT both suppressed genes involved in the inflammatory response and immune signaling pathways (**Figure 5.1F**), which are further repressed by the addition of Znf281(Zhou et al., 2017).

Focusing on the GO term "sarcomere organization" identified for GHMT>GMT/Control (C1), we observed a nearly uniform upregulation of the genes comprising this category (**Figures 5.1G** and **S5.1E**). For the "regulation of heart rate" group, we also observed strong upregulation of individual genes specifically in GHMT-transduced iPMs, although the entire group is not uniformly activated, as seen for the "sarcomere organization" group (**Figures 5.1G** and **S5.1E**). Importantly, we note that reprogramming with Hand2 alone did not significantly increase gene expression in either of these groups, suggesting that Hand2 is not sufficient to activate pacemaker- and sarcomere-related genes. Taken together, these results demonstrate that Hand2

uniquely promotes iPM formation with GMT by activating genes necessary for myocyte contraction and cardiac conduction.

Comparative RNA-Seq analysis defines shared markers of iPMs and SAN.

To independently evaluate the extent to which iPM transcriptional reprogramming resembles endogenous pacemaker cells, we sought to compare our RNA-seq data with previous datasets(Vedantham et al., 2015). In contrast to our RNA-seq data, which underwent standard library preparation, the prior datasets required extensive library amplification due to the limited amount of obtained tissue. Therefore, we were unable to make direct quantitative comparisons between the two datasets. Using an alternative analytical approach, we qualitatively compared genes enriched in GHMT- versus GMT-transduced fibroblasts with those enriched in endogenous Hcn4-GFP^{+} SAN versus neighboring atrial tissue obtained at E14.5, P4, and P14 (**Figure 5.2A**).

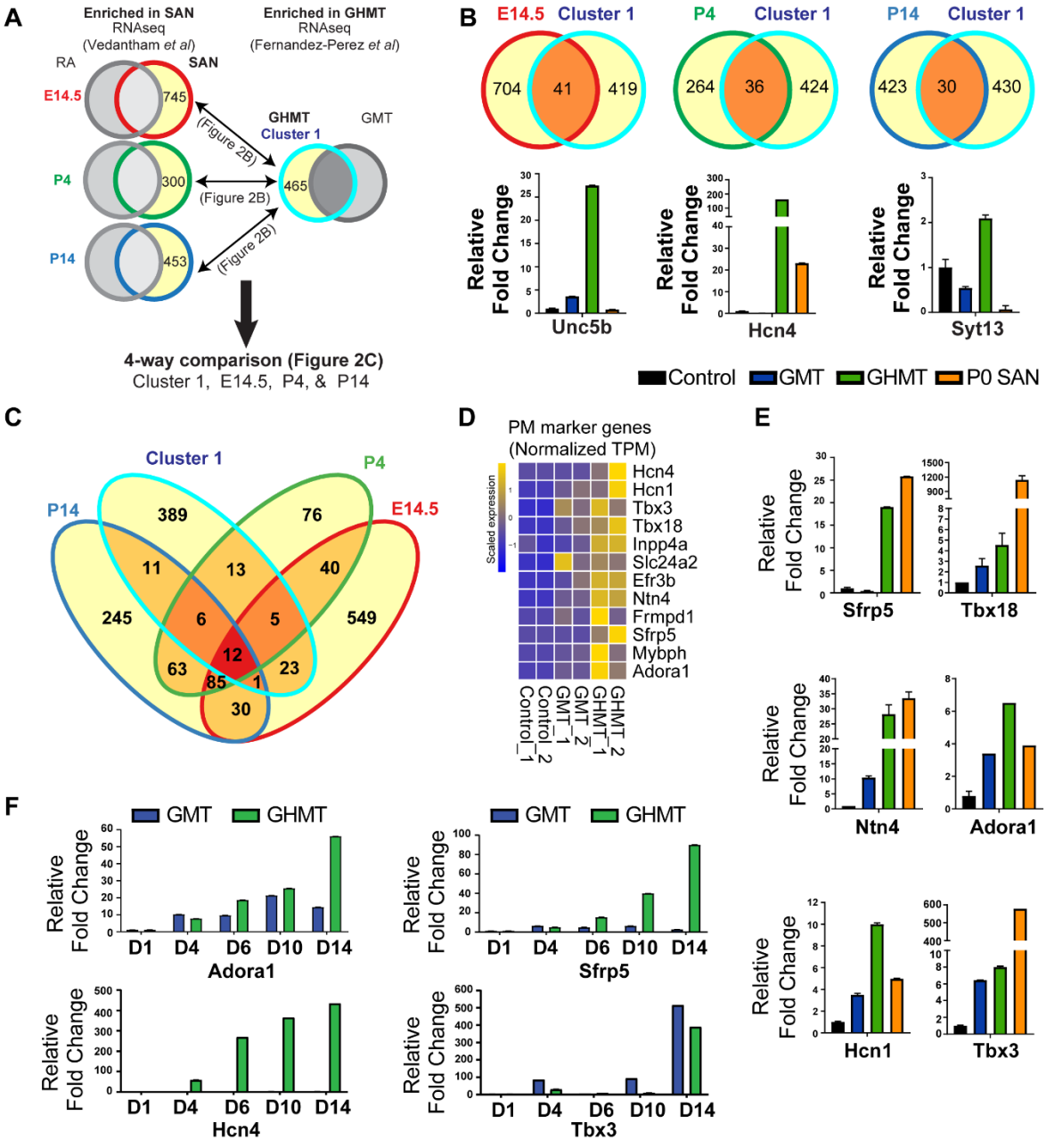


Figure 5.2. Core endogenous pacemaker marker genes are enriched in iPMs. *A) Analytical strategy for comparing iPM-specific genes with sinoatrial node (SAN)-specific genes in (B) and (C). Gene sets enriched in SAN compared to RA at 14.5, P4, and P14 (red outline with yellow interior) were obtained from Vedantham et al. Gene sets enriched in GHMT- versus GMT-reprogrammed cells (blue outline with yellow interior) were obtained from this study. Then, pairwise comparisons for overlapping genes at individual developmental time points were done in (B), while a 4-way comparison for overlapping genes amongst all gene sets was performed in (C). B) Top: Venn diagrams demonstrating numbers of shared genes between iPMs and each developmental time*

point analyzed in the previous study. Bottom: qRT-PCR for individual candidate genes are shown. C) Four-way Venn diagram shows a core set of 12 pacemaker marker genes. D) RNA-seq data for pacemaker marker genes across reprogrammed MEF samples. E) qRT-PCR validation of a subset of pacemaker marker genes with micro-dissected mouse SAN tissue as a positive control. F) qRT-PCR time course of GMT- and GHMT-transduced MEFs.

From individual pairwise comparisons, we observed 41, 36, and 30 overlapping genes between iPMs and SAN at E14.5, P4, and P14, respectively, and enrichment of individual genes was confirmed by qRT-PCR analysis (**Figure 5.2B**). Based on this qualitative comparison, we conclude that iPMs share a modest degree of gene signatures with endogenous SAN tissue across developmental time points. An integrated four-way comparison identified a core set of 12 shared iPM- and SAN-specific marker genes (**Figure 5.2C**). Visualization of the core PM marker genes in our RNA-seq data revealed highly specific expression in GHMT-transduced, but not uninfected or GMT-transduced, fibroblasts (**Figures 5.2D** and **S5.1E**).

A subset of PM marker genes was independently validated at the RNA level by qRT-PCR analysis (**Figure 5.2E**). In total, we performed qRT-PCR validation on 7 out 12 PM genes, 6 of which were successfully confirmed. Of the 7 core PM genes that were tested, qRT-PCR failed to robustly validate Tbx3 enrichment in iPMs. However, we emphasize that qRT-PCR validation was performed on unsorted cells, which may potentially dilute any signals of enrichment. Additional validation for Tbx3 at the protein level was obtained by immunocytochemistry (ICC) on GHMT-transduced fibroblasts (see **Figure 5.3F**). Interestingly, parallel experiments performed on GMT-reprogrammed

iCLMs failed to stain for Tbx3 (**Figure S5.1F**), suggesting that iPMs are uniquely positive for Tbx3 protein.

To investigate the kinetics of pacemaker marker genes during iPM formation, we performed time-course qRT-PCR on GMT- and GHMT-infected fibroblasts (**Figure 5.2F**). Interestingly, we observed distinct gene expression dynamics for each pacemaker marker gene during iPM formation. Collectively, these data support the idea that direct reprogramming is distinct from developmental fate specification. In this regard, iPMs do not share gene expression signatures with endogenous SAN tissue at any specific developmental time point, but rather seem to reflect a spectrum of developmental stages. Further evidence for this conclusion derives from the observation that the early SAN developmental transcription factors Tbx3 and Tbx18 are activated in iPMs, yet Shox2 and Isl1 are not (see **Figure 5.3E**). Our results also identify a novel set of shared iPM and SAN markers that are activated with varying kinetics.

Hand2 drives chromatin accessibility towards an endogenous pacemaker state.

Although transcriptional signatures drive lineage specification, changes in chromatin accessibility can illuminate mechanisms by which gene expression is modulated(Long et al., 2016). Therefore, we used the assay for transposon accessible chromatin sequencing (ATAC-seq) to comprehensively profile chromatin accessibility(Buenrostro et al., 2013; Corces et al., 2017). ATAC-seq libraries were generated from GFP-sorted control, GMT-infected, and GHMT-infected fibroblasts. Sorted endogenous P0 Hcn4-GFP^{+} pacemaker cells (P0 PMs) were used to generate

ATAC-seq libraries for comparison. Replicate ATAC-seq libraries showed a high degree of concordance, and the GMT and GHMT samples clustered together by similarity (**Figure S5.2A**). Unsupervised clustering analysis of chromatin accessibility peaks identified five unique clusters (1-5) representing distinct patterns of chromatin accessibility across Control, GMT, GHMT, and P0 PM samples (**Figures 5.3A** and **S5.2B**).

For subsequent analysis, we focused on the behavior of chromatin accessibility peaks that were divergent between the starting cell population (MEFs) and the target cell population (P0 PMs). Therefore, we explored clusters 3 (C3: Control>GMT/GHMT/P0 PM) and 5 (C5:P0 PM>GHMT>GMT>Control) in more detail. In aggregate, chromatin accessibility diminished for Control>GMT/GHMT/P0 PM (C3) and increased for P0 PM>GHMT>GMT>Control (C5) when comparing GHMT to GMT (**Figure S5.2B**), suggesting that Hand2 promotes epigenome remodeling towards a more PM-like state. Interestingly, P0 PM>GHMT>GMT>Control (C5) was relatively enriched for distal intergenic sites and depleted of proximal elements compared to Control>GMT/GHMT/P0 PM (C3) (**Figure S5.2C**).

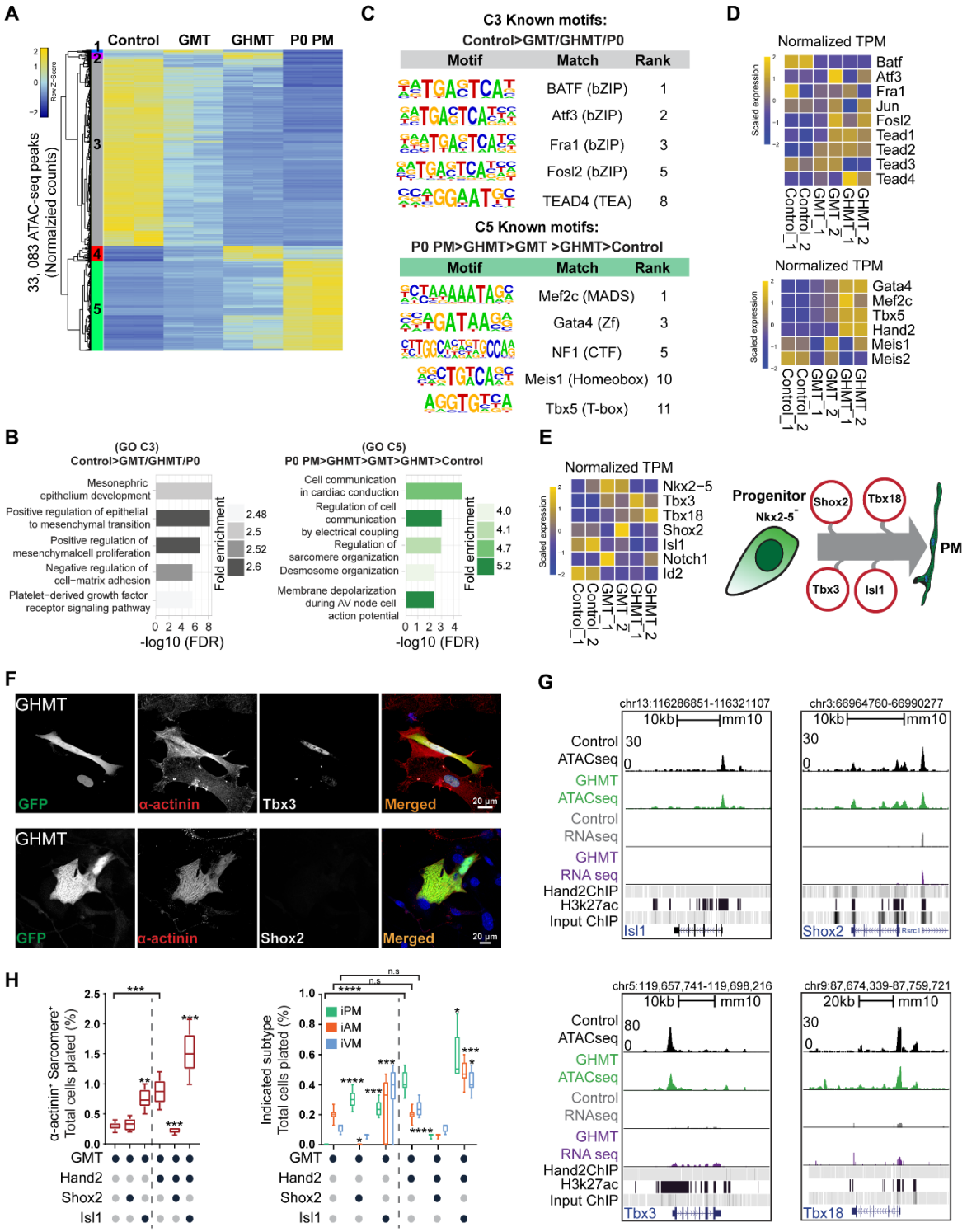


Figure 5.3. Hand2 reorganizes chromatin accessibility towards a pacemaker-like state. *A) ATAC-seq was performed on control, GMT-infected, and GHMT-infected MEFs, and P0 Hcn4-GFP⁺ endogenous pacemaker (P0 PM) cells were profiled for comparison. Unsupervised hierarchical clustering separated chromatin peaks into five clusters (1-5). B) Gene ontology (GO) term analysis of clusters 3 and 5. C) Known transcription factor motifs identified for clusters 3 and 5. D) RNA-seq enrichment of candidate transcription*

factors (TFs) for clusters 3 and 5. E) Cartoon diagram of the developmental pacemaker gene regulatory network (right) and associated RNA-seq data (left). F) Immunocytochemistry (ICC) on GHMT-reprogrammed cells for Tbx3 and Shox2. Scale bar, 20μm. G) Genome browser tracks for the Isl1, Shox2, Tbx3, and Tbx18 loci with associated ATAC-seq and RNA-seq tracks. Cardiac Hand2 ChIP-seq and H3K27Ac ChIP-seq peaks are shown as a reference. H) Cardiac reprogramming of MEFS was assessed for sarcomere formation (left) and subtype specification (right) in the presence of the indicated PM gene regulatory TFs (mean±SD, n=6).

Next, we performed GO term analysis using the nearest annotated neighboring genes for individual chromatin accessibility peaks (**Figure 5.3B** and **S5.2D**). We found that the terms associated with Control>GMT/GHMT/P0 PM (C3) chromatin included alternative fates ("mesonephric epithelium development"), fibroblast characteristics ("positive regulation of mesenchymal cell proliferation"), and signaling pathways ("platelet-derived growth factor receptor signaling pathway"). In contrast, terms associated with P0 PM>GHMT>GMT>Control (C5) chromatin were similar to GO terms for iPM gene expression signatures (**Figure 5.1F**) and included the terms "regulation of cell communication by electrical coupling," "regulation of sarcomere organization," and "membrane depolarization during AV node cell action potential." Altogether, these results strongly suggest that iPM chromatin accessibility mirrors observed changes in gene expression.

Changes in chromatin accessibility are typically orchestrated by coactivators, corepressors, and chromatin remodelers that are recruited to genomic DNA by sequence-specific TFs(Long et al., 2016). Therefore, we searched Control>GMT/GHMT/P0 PM (C3) and P0 PM>GHMT>GMT>Control (C5) chromatin accessibility peaks for recurrent motifs that bind known TFs (**Figures 5.3C** and **S5.2E**). For Control>GMT/GHMT/P0 PM (C3) peaks, we observed a preponderance of binding motifs for the bZIP class of TFs,

including Atf3, Fra1, and AP-1 (Fos/Jun). bZIP TFs play important roles in fibroblast gene expression and their binding motifs are over-represented in closing chromatin during iPS reprogramming(Li et al., 2017). In contrast, we identified Mef2c, Gata4, and Tbx5 motifs in P0 PM>GHMT>GMT>Control (C5) chromatin, as would be expected for fibroblasts transduced with GHMT. Surprisingly, the Hand2 consensus motif was not detected in this analysis (see **Figure 5.4** and Discussion). Intriguingly, however, our analysis did identify motifs for TEAD4 and Meis1, both of which have been implicated in heart development(Chen et al., 1994; Stankunas et al., 2008).

The close alignment of the GO terms and motifs enriched in iPMs and endogenous pacemaker cells was particularly impressive for the P0 PM>GHMT>GMT>Control (C5) cluster. However, we wondered whether this was driven by the comparison with target cells or if iPMs genuinely resembled endogenous pacemaker cells. To distinguish between these two possibilities, we re-analyzed the Control, GMT, and GHMT datasets without the P0 PM samples (**Figure S5.3**). Unsupervised hierarchical clustering identified 5 clusters (**Figure S5.3A-B**). For Cluster 3 (GHMT>GMT>Control), which resembled Cluster 5 (P0 PM>GHMT>GMT>Control) in **Figure 5.3**, we note that GO terms related to cardiac conduction, such as “Regulation of heart rate,” “Regulation of heart rate by cardiac conduction,” and “cardiac conduction” were amongst the most enriched terms (**Figure S5.3C**). Furthermore, the list of motifs discovered in Cluster 3 (GHMT>GMT>Control) was highly concordant with the motifs identified for Cluster 5 (P0 PM>GHMT>GMT>Control) in **Figure 5.3** (**Figure S5.3D**). Taken together, these

observations strongly suggest that the alignment between iPM and P0 PM chromatin accessibility is not merely an artifact of our analytical method.

Given that specific binding motifs emerged from our ATAC-seq data, we next wished to evaluate whether the cognate TFs were expressed and/or enriched in GHMT-transduced iPMs. Therefore, we interrogated our companion RNA-seq datasets for TFs whose motifs were enriched in P0 PM>GHMT>GMT>Control (C5) or Control>GMT/GHMT/P0 PM (C3) chromatin (**Figures 5.3D** and **S5.4A**). For the motifs identified in Control>GMT/GHMT/P0 PM (C3) chromatin, we observed that bZIP TFs were not specifically enriched or depleted in iPMs, while Tead1, Tead2, and Tead4 showed increased expression in GHMT-transduced fibroblasts compared with controls. For motifs obtained from P0 PM>GHMT>GMT>Control (C5) chromatin, we reassuringly found increased expression of Gata4, Mef2c, Tbx5, and Hand2 in iPMs, while expression of Meis1/2 was less consistent. Collectively, these results demonstrate that global chromatin accessibility in GHMT-transduced iPMs moves towards a PM-like state. However, it remains unclear whether iPM reprogramming recapitulates key aspects of PM development.

Based on established gene regulatory networks that drive developmental specification of pacemaker cells(van Eif et al., 2018), we were particularly interested in assessing whether these networks were re-deployed during iPM reprogramming (**Figures 5.3E** and **S5.4A**). Interestingly, we observed that Tbx3 and Tbx18 were upregulated in iPMs, but Shox2 and Isl1 were not (**Figures 5.3E** and **S5.4A**). Expression

of Notch1 and Id2, which regulate distinct aspects of cardiac conduction system development(van Eif et al., 2018), was not increased in iPMs, and expression of Nkx2.5, a negative regulator of pacemaker specification, was diminished (**Figures 5.3E** and **S5.4A**).

Expression levels for Tbx3, Tbx18, Shox2, and Isl1 were independently confirmed by qRT-PCR on separate samples (**Figures 2E** and **S4B**). Furthermore, protein levels of endogenous Tbx3 and Shox2 in iPMs were confirmed by ICC (**Figure 5.3F**), which was consistent with the RNA-seq and qRT-PCR results. Interestingly, we found little correlation between chromatin accessibility and gene expression for the PM regulatory TFs (**Figure 5.3G**). For example, chromatin was relatively accessible for all loci in fibroblasts and changed minimally in iPMs, yet only Tbx3 and Tbx18 expression was detectable in iPMs (**Figures 5.3E** and 5**.3G**). Together, these observations suggest that substantial chromatin reorganization at developmentally important loci is not required for transcriptional activation during iPM reprogramming.

Analysis of the iPM ATAC-seq and RNA-seq datasets thus far demonstrated the following: 1) Shox2 and Isl1 are inactive members of the developmental pacemaker gene regulatory network, 2) opening loci are enriched for Meis1 and GMT motifs, and 3) closing loci are enriched for bZIP motifs. Based on these findings, we hypothesized that the addition of TFs from each of these categories would perturb iPM reprogramming either positively (Meis1/Isl1/Shox2) or negatively (bZIP). To test these possibilities, we

performed iPM reprogramming with GMT and GHMT with or without single candidate factors.

When Tbx3 or Tbx18 were added to GMT or GHMT, we observed no detectable increase in either sarcomere$^+$ iCLMs or iPMs (**Figure S5.4C**), which is consistent with the fact that GHMT already activates endogenous Tbx3 and Tbx18. In contrast, Isl1 and Shox2 each had interesting effects on cardiac reprogramming. Isl1 substituted for Hand2 to drive both sarcomere$^+$ iCLMs and iPMs, while Isl1 added to GHMT further improved both iCLM and iPM reprogramming (**Figure 5.3H**). Shox2 had more subtle effects. When Shox2 was added in place of Hand2, overall sarcomere$^+$ iCLM production was unaffected, but the resulting iCLMs were skewed towards iPMs (**Figure 5.3H**). In contrast, both iCLM and iPM generation were reduced when Shox2 was added to GHMT (**Figure 5.3H**). The addition of Meis1 or Atf3 to GHMT did not affect sarcomere$^+$ iCLMs, but we observed a shift from iVMs to iAMs for Meis1 and suppression of iPMs by Atf3 (**Figure S5.4C**). We note that addition of Atf3 or Meis1 uniquely generated mixed iCLMs (Hcn4-GFP$^+$/Myl2$^+$), which we have not observed previously (**Figure S5.4D**). Interestingly, the addition of Meis1 to GMT was capable of generating iPM (**Figure S5.4B**). Taken together, our data demonstrate coordinated chromatin accessibility changes in GHMT-transduced iPMs towards an endogenous pacemaker-like state. Focusing on established PM gene regulatory networks, our results identify Isl1, Shox2, and Meis1 addition as potential routes for rational improvement of iPM formation by GHMT.

Identification of Hand2-dependent iPM cis-regulatory elements.

Given that iPM reprogramming is dependent upon Hand2, we sought to determine how chromatin accessibility changes relative to established Hand2 genomic binding sites. To address this question, we took advantage of a previously generated Hand2 ChIP-seq dataset derived from mouse embryonic heart(Laurent et al., 2017). Importantly, de novo motif discovery on this Hand2 ChIP-seq dataset showed that the consensus CATCTG Hand2 binding site was amongst the most common E-box motifs(Dai and Cserjesi, 2002; Laurent et al., 2017). We intersected this dataset (n=12,101 peaks) with open iPM ATAC-seq peaks (n=33,032 peaks) to identify a set of 1,684 Hand2 genomic binding sites that are accessible in iPMs (**Figure 5.4A**). Using this defined set of known Hand2 bindings sites, which we refer to as the iPM Hand2 cistrome, unbiased clustering was performed across all samples, including fibroblasts infected with Hand2 alone (**Figure 5.4A**). Interestingly, this analysis separated the iPM Hand2 cistrome into 4 distinct clusters. Clusters 1, 3, and 4 were accessible in both iPMs and P0 PMs, while cluster 2 had limited accessibility in P0 PMs.

We mapped the genomic location of individual clusters to gain a better appreciation for unique features of each group (**Figure S5.5A**). This analysis revealed that the vast majority of Hand2 binding sites occurred at distal intergenic or intronic regulatory elements. Nevertheless, there were subtle differences between each cluster of Hand2 binding sites. For example, cluster 3 had a higher proportion of promoter-proximal Hand2 binding sites, while cluster 2 tended to have binding sites in the first intron (**Figure S5.5A**). GO terms for each iPM Hand2 cistrome cluster returned revealing gene categories (**Figures 5.4B** and **S5.5B**). The closing Hand2 genomic binding sites in cluster 2 (C2:

Control/Hand2>GMT/GHMT>P0 PM) associated with generic terms related to developmental morphogenesis and actin filaments (**Figure 5.4B**). In contrast, the opening Hand2 genomic binding sites in cluster 3 (C3: P0 PM/GHMT>Control/Hand2) returned terms related to the cardiac conduction system, including “cell-cell signaling involved in CCS” and “Bundle of His development” (**Figure 5.4B**).

Next, we searched for recurrent TF co-binding motifs in each cluster of Hand2 genomic binding sites (**Figure S5.5C**). As expected, motifs for GATA, MEF2, and T-box families of TFs were prevalent in the opening clusters (1, 3, and 4), while bZIP and TEAD motifs were identified in the closing cluster 2 (**Figure S5.5C**). Since Gata4, Hand2, Mef2c, and Tbx5 are known to function cooperatively(Dai and Cserjesi, 2002; Garg et al., 2003; Morin et al., 2000; Zang et al., 2004), we identified genomic locations for GATA, MEF2, and T-box family binding motifs to assess potential co-occupancy of GHMT across the iPM Hand2 cistrome (**Figure 5.4C**). Interestingly, we found a large number of potential Hand2-Gata4-Mef2c (H-G-M), Hand2-Mef2c-Tbx5 (H-M-T), and Hand2-Gata4-Tbx5 (H-G-T) composite binding sites dispersed across the iPM Hand2 cistrome. Given the flexibility of the E box motif (CANNTG), we also tested the hypothesis that subtle alterations in the Hand2 binding motif may account for the distinct chromatin accessibility profiles of clusters 1-4. However, we found no evidence that specific E boxes segregated with a particular cluster (**Figure S5.5D**). Taken together, these observations suggest that GHMT cooperate to activate the iPM Hand2 cistrome.

Within the iPM Hand2 cistrome, we noticed that chromatin accessibility at the Hcn4 locus was dramatically increased in GHMT-reprogrammed iPMs, but not in the presence of either GMT or Hand2 alone (**Figure 5.4D**). This observation suggests that the Hcn4 locus undergoes dramatic chromatin reorganization during iPM reprogramming only when Hand2 is transduced in the presence of GMT. Specifically, we identified four ATAC-seq peaks that coincided with both Hand2 ChIP-seq peaks and H3K27ac ChIP-seq peaks from embryonic heart. Thus, we nominated these four putative cis-regulatory elements as candidate Hand2-dependent Hcn4 regulatory elements. To test their functionality, we cloned each element upstream of a minimal TATA box and a luciferase transgene. Using these reporter constructs, we performed transient transfection analysis in HEK cells in the presence of various TF constructs (**Figure 5.4E**).

Given that a consensus Hand2 binding motif did not emerge from our analysis of accessible chromatin in iPMs (**Figures 5.3C, S5.3D,** and **S5.5C-D**), we hypothesized that Hand2 would only activate Hcn4 transcription cooperatively with GMT. In support of this idea, Hand2 did not significantly activate any of the Hcn4 elements with the exception of a mild effect on the Region 1 reporter (**Figure 5.4E**). Importantly, this lack of transcriptional activation was not a general property of E-box binding transcription factors, since the effect of MyoD on each of these reporters diverged from the response to Hand2 for the Region 1 and 4 reporters (**Figure 5.4E**). Consistent with the notion that Hand2 activates Hcn4 transcription in conjunction with GMT, we observed that GHMT activated the Region 2 and 4 reporters (**Figure5. 4E**). In contrast, GHMT did not activate Region 1 and 3 reporters (**Figure 5.4E**). To exclude the possibility that these discordant

observations were due to the inability of Hand2 to bind Region 1, we performed parallel transfection studies using a Hand2-VP16 construct to demonstrate activation of both Region 1 and 2 reporters (**Figure S5.5E**), suggesting that Hand2 is capable of binding both regulatory elements as shown by ChIP-seq(Laurent et al., 2017). Taken together, these results confirm that Hcn4 Regions 1-4 are likely to be Hand2 response elements yet hint at functional nuances for each regulatory element.

The Hand2 N-terminus harbors a cardiac subtype diversification (CSD) domain.

To explore the biochemical mechanisms by which Hand2 regulates iPM formation, we performed extensive structure-function analysis. To address whether Hand2 functions as an activator or repressor during cardiac reprogramming, we created Hand2-VP16 and Hand2-EnR fusion proteins and tested their ability to support cardiac reprogramming with GMT (**Figure S5.6A-B**). We found that Hand2-VP16 slightly improved and Hand2-EnR strongly repressed cardiac reprogramming, suggesting that Hand2 functions as an activator in this context. Using a previously established dimerization-defective Hand2 mutant ($Hand2^{F119P}$)(McFadden et al., 2002), we also demonstrated that Hand2 dimerization is required for optimal cardiac reprogramming (**Figure S5.6C**). Taken together, these data show that Hand2 functions as a dimeric transcriptional activator during cardiac reprogramming.

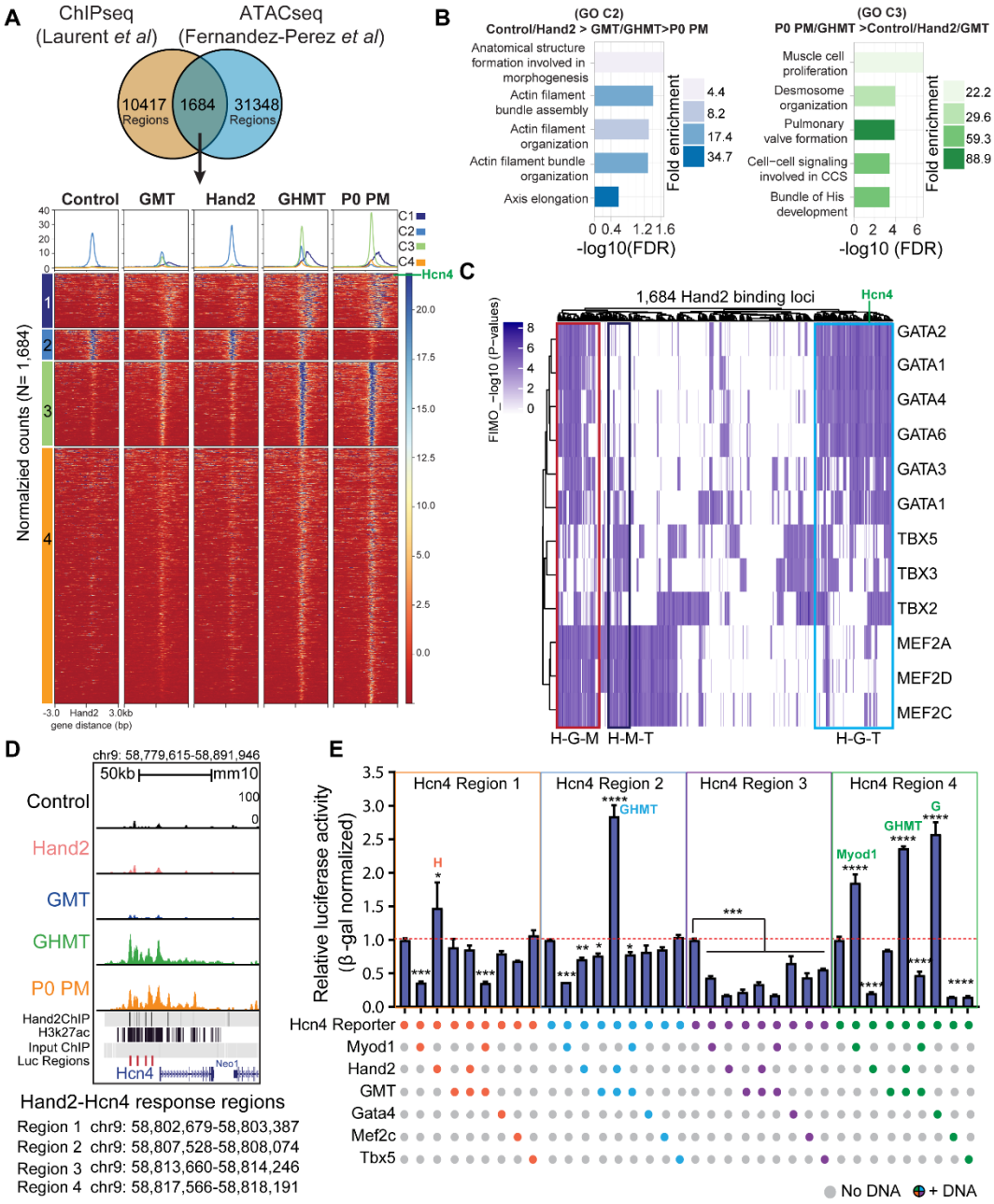


Figure 5.4. Modulation of Hand2 cistrome accessibility contributes to iPM reprogramming. *A) Cardiac Hand2 ChIP-seq and iPM ATAC-seq peaks were intersected to establish a set of 1684 iPM Hand2 cistrome peaks, which clustered into four groups (1-4). B) Gene ontology (GO) term analysis for clusters 2 and 3. C) Distribution of GMT binding motifs at 1684 Hand2 genomic binding sites accessible in iPMs. Putative composite sites are boxed and labeled. D) Genome browser track of the Hcn4 locus, demonstrating increased accessibility in the promoter-proximal region that coincides with H3K27Ac and Hand2 ChIP-seq peaks. The regions selected for functional validation are shown as red tick marks, and the corresponding genomic coordinates are given below. E) Transient transfection assays were performed in HEK cells with specific Hcn4 regulatory element reporters and the indicated expression constructs. Transfection*

*experiments were performed three times in triplicate (n=9), and a representative result is shown. Statistical analysis was performed using ANOVA with Bonferroni post-hoc correction for each set of samples transfected with the same reporter, and significant comparisons are noted. *, p<0.05; **, p<0.01; ***, p<0.001; ****, p<0.0001.*

The Hand2 protein sequence can be partitioned into unstructured N- and C-termini and a central bHLH domain, which regulates DNA binding, dimerization, and protein-protein interactions(Conway et al., 2010). To analyze the function of each domain, we performed cardiac reprogramming experiments in which fibroblasts were infected with GMT plus Hand2 or one of the variants shown in **Figure 5.5A**. Based on the presumption that Hand2 DNA-binding is required for efficient cardiac reprogramming, we were encouraged to find that transduction of the Hand2^{EDE} mutant, which cannot bind DNA(McFadden et al., 2002), led to reduced iPM formation (**Figure 5.5B**). Interestingly, neither the N-terminal nor C-terminal deletion Hand2 mutants could support optimal cardiac reprogramming with GMT, suggesting that each domain is functionally required.

Since Hand1 and Hand2 exhibit complementary expression patterns and a high degree of amino acid conservation (**Figure 5.5C**), they may function similarly within distinct anatomical compartments. However, gene replacement studies and in vitro binding site selection assays favor the notion that each is functionally distinct(Dai and Cserjesi, 2002; Firulli et al., 2010; Hollenberg et al., 1995). To test these alternatives in the context of cardiac reprogramming, we compared the outcome of infecting fibroblasts with GMT + Hand2 versus GMT + Hand1. Interestingly, we found that Hand1 could not support optimal cardiac reprogramming as compared to Hand2 (**Figure 5.5C**). Thus, we conclude that Hand2 uniquely augments cardiac reprogramming in the context of GMT.

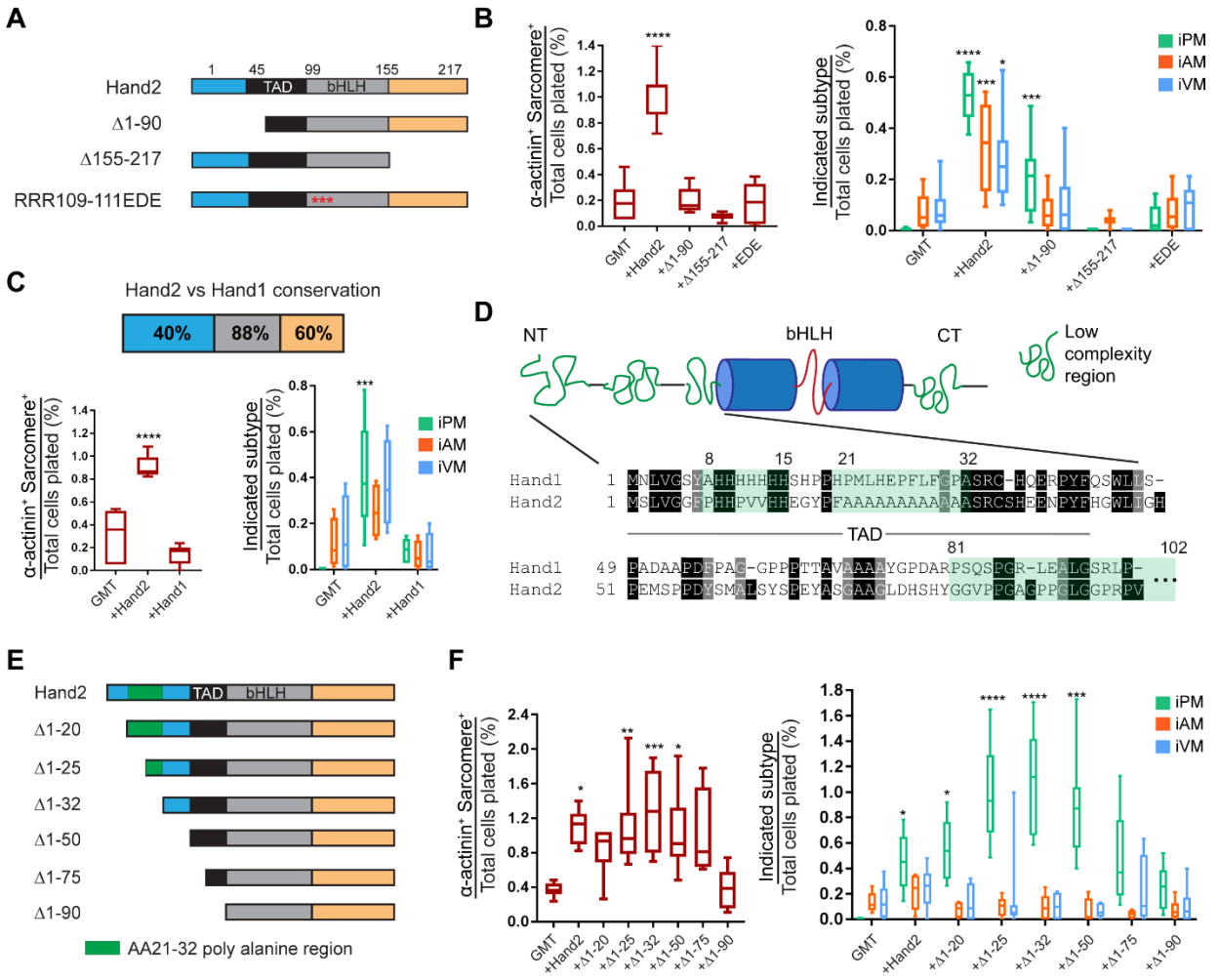


Figure 5.5. The Hand2 N-terminus encodes a subtype diversification domain (SDD). *A) Schematic for the Hand2 mutants used in (B). Unstructured N- and C-termini are shown in blue and orange, respectively, while known transactivation domain (TAD) is shown in black. Beta helix-loop-helix (bHLH) domain is shown in gray, and mutations in DNA-contacting basic residues in the RRR109-111EDE construct are depicted as red asterisks. B) Cardiac reprograming with GMT alone or with the indicated Hand2 mutants was assessed for sarcomere organization (left) and subtype diversification (right) (mean±SD, n=8). C) Cardiac reprogramming was conducted with GMT plus Hand2 or Hand1, and iCLMs were assessed for sarcomere organization (left) and subtype diversity (right). The degree of primary amino acid conservation between Hand2 and Hand1 for each protein domain is indicated at the top (mean±SD, n=6). D) Cartoon diagram of Hand2 predicted secondary structure with low complexity regions (LCRs) highlighted in green and bHLH region in blue cylinders and red line (top). The primary sequence of the N-terminus is shown below with the indicated LCRs. E) Schematic for the Hand2 N-terminal mutants used for reprogramming in (F). F) Cardiac reprogramming was performed with GMT alone or combined with the indicated Hand2 N-terminal mutants. iCLMs were assessed for sarcomere organization (left) and subtype diversity (right) (mean±SD, n=6).*

Although each domain of Hand2 appears to be functionally important, we chose to focus on the N-terminus for four reasons. First, the N- and C-termini remain functionally underexplored. Second, amino acid conservation between Hand1 and Hand2 is lowest for the N-terminus (**Figure 5.5C**), implying that differences in this region are likely to account for functional disparities between the two proteins. Third, a parallel experiment in which a VP16 activation domain was fused to Hand2Δ1-90 failed to restore cardiac reprogramming efficiency in combination with GMT (data not shown), suggesting that the N-terminus contains more than just the previously established transactivation domain(Dai and Cserjesi, 2002). Fourth, the Hand2 N-terminus harbors three predicted low complexity regions, including a stretch of poly-alanines (**Figure 5.5D**), which have been implicated in several human diseases associated with TF gene mutations(Albrecht and Mundlos, 2005).

To rigorously evaluate the function of the Hand2 N-terminus, we constructed a series of N-terminal deletion mutants (**Figure 5.5E**) and subjected them to several quality-control measures. ICC of infected MEFs confirmed nuclear localization of all Hand2 constructs (**Figure S5.6D**). Western blotting demonstrated production of full-length Hand2 proteins (**Figure S5.6E**). Transient transfection assays with an αMHC reporter construct and Hand2 mutant VP16 fusions suggested intact DNA binding ability (**Figure S5.6F**). Finally, co-immunoprecipitation assays confirmed that Hand2 mutants retained the ability to interact with known binding partners (**Figure S5.6G**). Taken together, these results demonstrate that Hand2 N-terminal mutants maintain proper nuclear localization, protein expression, DNA binding, and protein-protein interactions.

Using the collection of Hand2 N-terminal mutants, we systematically tested their ability to mediate cardiac reprogramming of MEFs along with GMT (**Figure 5.5F**). These studies led to two interesting observations. First, we found that truncation to amino acid 75 had no major effect on overall cardiac reprogramming, as the number of sarcomere^{+} cells remained unaffected (**Figure 5.5F**). Consistent with the presence of the previously described transactivation domain between amino acids 50 and 90, however, we observed that cardiac reprogramming diminished substantially when the Hand2Δ1-90 mutant was used (**Figure 5.5F**). Second, we noted profound changes in the proportion of iPMs generated by specific Hand2 N-terminal deletion mutants. In particular, Hand2 mutants in which the transactivation domain remained intact, yet any portion of the poly-alanine stretch was compromised (e.g. Hand2Δ1-25, Hand2Δ1-32, and Hand2Δ1-50), led to dramatically increased production of iPMs at the expense of iAMs and iVMs (**Figure 5.5F**). The iPMs generated by the poly-alanine mutant Hand2 constructs appeared morphologically indistinguishable from those generated by GHMT (**Figure S6H**). Based on these experiments, we conclude that the Hand2 N-terminus harbors both a transactivation domain and a novel cardiac subtype diversification (CSD) domain.

Integrated analysis uncovers desmosome gene activation in iPMs.

Encouraged by the insights provided by our RNA-seq and ATAC-seq data analysis, we sought to identify potential milestones on the path towards generating pacemaker cells. Therefore, we performed integrated analysis of the differences between GHMT- and GMT-transduced fibroblasts to illuminate Hand2-dependent gene regulatory networks that reflect acquisition of a PM-like transcriptional state. To accomplish this

objective, we simultaneously compared ATAC-seq and RNA-seq datasets derived from GHMT- and GMT-infected fibroblasts. We plotted RNA-seq and ATAC-seq peaks as a fold change between the GHMT and GMT samples. In this manner, peaks that appear in the right upper quadrant represent genes whose loci become more accessible and whose expression increases. Likewise, peaks in the left lower quadrant represent genes whose loci become less accessible and whose expression decreases. Confirming the overall feasibility of this approach, we found that genes comprising the "sarcomeric organization" GO term (**Figure 5.1G**) and our newly identified pacemaker marker genes (**Figure 5.2D**) tended to appear in the right upper quadrant (**Figure S5.7A**).

Focusing on genes in the right upper quadrant, we identified a group of desmosomal genes (**Figure 5.6A**). We observed concerted upregulation of the desmosomal gene battery specifically in iPMs (**Figures 5.6B** and **S5.7B**). Desmosomes are cellular structures that allow tight connections between cells and consist of several individual protein components (**Figure 5.6C**). Desmosomes are critical for intercellular communication, and human mutations in desmosomal genes can cause familial arrhythmia syndromes(Corrado et al., 2017). Furthermore, mouse knockout studies have clearly demonstrated that the desmosomal protein desmoplakin is required for normal sinus node function(Mezzano et al., 2016).

Consistent with the integrated analysis, we observed that chromatin accessibility at the Dsp1 locus increased substantially between GMT and GHMT samples, whereas accessibility at the Pkp2 locus increased less dramatically (**Figure 5.6D**). Interestingly,

we note that multiple Hand2 ChIP-seq peaks underlie regions of increased chromatin accessibility, suggesting that Hand2 is a primary driver of this process. To confirm the RNA-seq data, we performed qRT-PCR analysis on a separate set of reprogrammed fibroblasts (**Figure 5.6E**). As expected, we found increased expression of desmosomal components Dsp, Pkp2, Dsc2, Jup, and Dsg2 in GHMT- versus GMT-infected fibroblasts. Moreover, we observed that endogenous Pkp2 was expressed at the protein level in GHMT-infected, but not GMT-infected, iCLMs (**Figure 5.6F**).

To evaluate desmosome gene expression dynamics, we performed qRT-PCR on GHMT-infected fibroblasts at various time points following transduction (**Figure 5.6G**). Although all tested genes increased expression after GHMT transduction, each desmosomal component exhibited unique expression dynamics and maximal levels. A recent study suggested that Pkp2 can modulate gene expression(Cerrone et al., 2017), so we tested this possibility in the context of iPM reprogramming (**Figure S5.7C**). However, we found no evidence that Pkp2 could either substitute for Hand2 or augment GHMT to improve iPM formation. Taken together, these results identify a Hand2-dependent desmosome gene battery that undergoes coordinated chromatin reorganization and gene activation during iPM reprogramming.

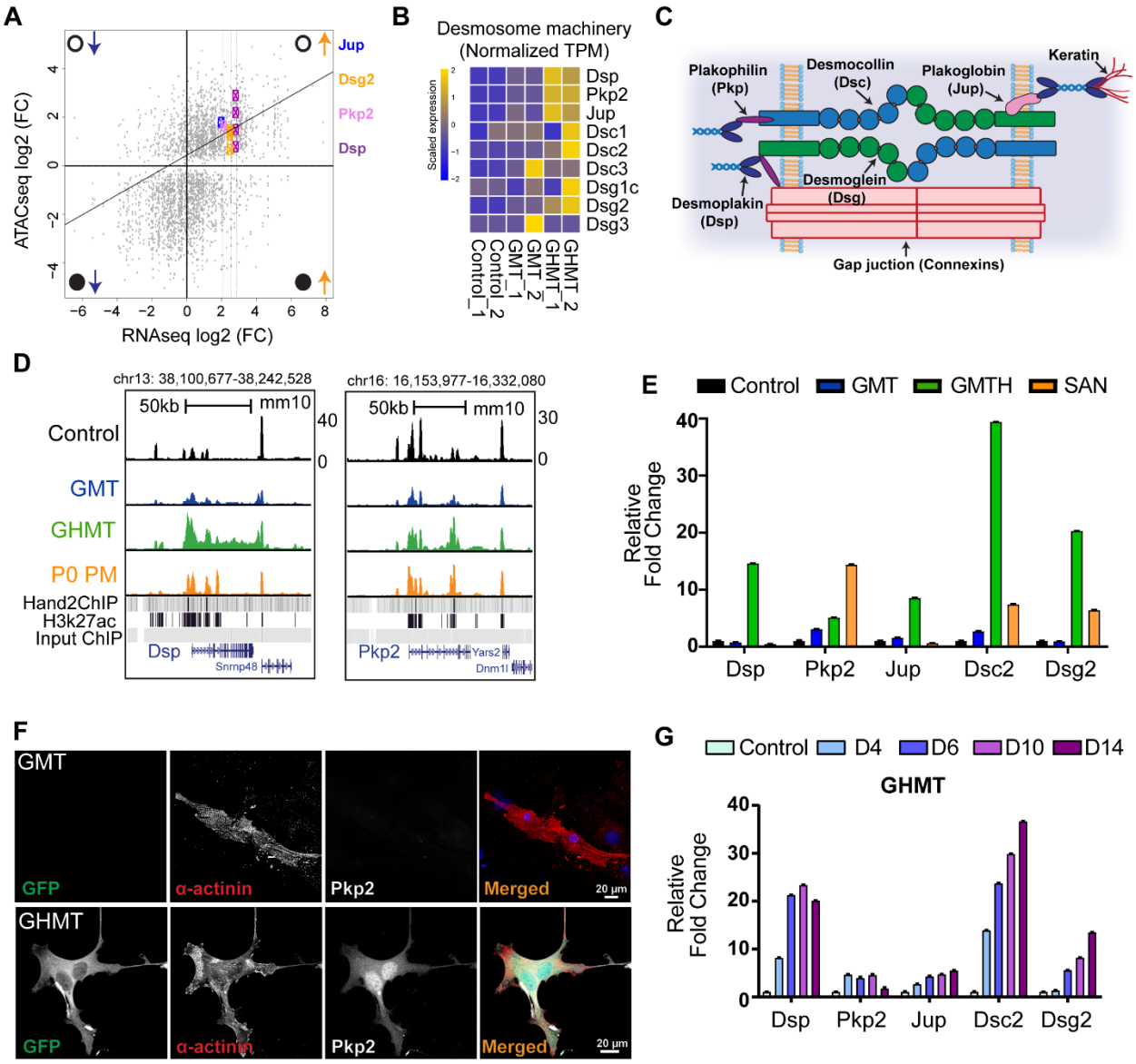


Figure 5.6. Hand2 coordinates activation of the desmosome machinery. *A) Integrated analysis of RNA-seq (X-axis) and ATAC-seq (Y-axis) data between GHMT- and GMT-transduced MEFs. Open and closed circles denote open and closed chromatin, while upward and downward pointing arrows indicated increased and decreased gene expression. Gene loci are labeled with colored symbols and corresponding gene names. B) RNA-seq heatmap for genes that comprise the desmosome machinery in control, GMT-infected, and GHMT-infected MEFs. C) Cartoon diagram of the desmosome complex. D) Genome browser tracks Dsp and Pkp2 with corresponding ATAC-seq tracks derived from GMT, GHMT, and P0 PM samples. Hand2 and H3K27 ChIP-seq peaks are provided as a reference. E) qRT-PCR confirmation that desmosome genes are enriched in iPMs. F) ICC validation of Pkp2 protein expression in GMT- and GHMT-transduced MEFs. Scale bar, 20μm. G) Kinetics of desmosome gene expression during iPM formation demonstrated by qRT-PCR analysis.*

DISCUSSION

In this study, we conducted a detailed genomic and biochemical analysis of the mechanisms by which Hand2, in cooperation with Gata4, Mef2c, and Tbx5, coordinates iPM formation. Hand2 orchestrates the activation of a panel of endogenous pacemaker marker genes that span embryonic and postnatal SAN development. Interestingly, we found that Shox2 and Isl1, which participate in the developmental pacemaker gene regulatory network, are not activated in iPMs, and they can augment iPM formation to some extent when added exogenously. Furthermore, we demonstrated that Hand2 cistrome accessibility undergoes coordinated changes to open loci of cardiac conduction genes, including Hcn4. Through structure-function analysis of Hand2, we discovered a novel cardiac subtype diversification domain (CSD) residing within the unstructured N-terminus. Finally, we found that a battery of desmosomal genes are specifically activated in the presence of Hand2 via concerted chromatin reorganization.

To gain a sense for the degree of similarity observed between iPMs and SAN (**Figure 2**), we note that in the original comparison of SAN and RA RNA-seq datasets(Vedantham et al., 2015), there are 142, 166, and 128 overlapping genes between the SAN-enriched gene list for E14.5 vs. P4, P4 vs. P14, and E14.5 vs. P14, respectively. Thus, we consider the degree of overlap between iPMs and SAN tissue (**Figure 5.2B**) to be at least moderate. Furthermore, we believe that the degree of observed overlap likely represents an underestimate for the following three reasons. First, the RNA amplification required for the SAN-enriched dataset(Vedantham et al., 2015) would lead to overrepresentation of certain genes and skewing of the comparison

with the iPM dataset, which was generated without amplification. Second, the SAN-enriched RNA-seq dataset was generated by laser capture such that both CMs and non-CMs (as well as some Hcn4-GFP^{-}) tissue may have been obtained, thus leading to a more heterogeneous sample composition and a dilution of the similarity with our iPM dataset. Third, iPMs are engineered cells that approach a transcriptional signature of SAN tissue, yet they remain short of the target population.

Since Isl1 is known to regulate pacemaker formation (Liang et al., 2015), we were surprised that Isl1 transcripts were not increased in GHMT-reprogrammed cells (**Figure 5.3E).** We surmised that the Isl1 transcriptional network failed to be engaged by GHMT, which motivated our decision to test it as a candidate reprogramming factor. Consistent with this hypothesis, we found that addition of Isl1 to GHMT improved iPM reprogramming, but Isl1 could also partially substitute for Hand2 (**Figure 5.3H**). Interestingly, the Isl1 motif did appear in the HOMER analysis of cluster 5, but it was only enriched 1.1-fold above background, so we did not consider this level of enrichment meaningful. To reconcile these observations, we posit that Hand2 and Isl1 activate a set of common transcriptional targets requiring either TF alone and a separate set of targets requiring both. Presumably, the latter category of target genes involve cooperative binding of a Hand2-Isl1 complex. This proposed mechanistic model is consistent with either direct or indirect gene regulation, and distinguishing between them represents an important area of future investigation.

Surprisingly, the consensus Hand2 motif (CATCTG) did not emerge from our HOMER analysis of the iPM ATAC-seq data (**Figure 5.3C**). However, careful review of the raw analysis (Tables S4-S8) reveals motifs for other bHLH family members in clusters 3-5 using less stringent criteria for motif discovery. Overall, none of the identified bHLH motifs enrich substantially above background, suggesting that Hand2 does not bind to a strong E-box consensus sequence during iPM reprogramming. Furthermore, a strong E-box motif did not emerge from our analysis of iPM ATAC-seq peaks that overlap with known Hand2 binding sites as determined by ChIP-seq (**Figure S5.5C-**D). We believe that context-dependent alterations in Hand2 DNA binding affinity helps to explain these seemingly contradictory observations.

The original Hand2 consensus motif derived from in vitro binding site selection assays is CATCTG(G) with very little flexibility(Dai and Cserjesi, 2002). Interestingly, the derived Hand2 consensus based on the embryonic limb Hand2 ChIP-seq dataset is CRTCTG(G) with “A” being the preferred nucleotide at the “R” position(Osterwalder et al., 2014). This consensus is almost identical to the one identified in vitro, suggesting that Hand2 by itself is capable of binding to these motifs in the limb. In contrast, the Hand2 consensus based on the embryonic heart Hand2 ChIP-seq dataset is CAKCTG(B)(Laurent et al., 2017). Here, the “K” and “B” positions are evenly divided amongst all nucleotide possibilities and are not strongly specified overall, suggesting that the consensus is much more flexible in the heart. We believe that this reflects different binding modes in limb versus heart and iPMs, which require cooperative binding of Hand2 with GMT. Consistent with this idea, a Hand2 DNA-binding deficient mutant can rescue

heart, but not limb, development(Liu et al., 2009), although we note that this mutant cannot completely rescue efficient iPM formation (**Figure 5.5B**).

Given that Hand2 coordinates activation of cardiac conduction genes, we hypothesize that Hand2 may also regulate developmental pacemaker specification. Two observations provide intriguing support for this idea. First, Hand2 is unique amongst GHMT to activate pacemaker genes. Second, several known cardiac genomic Hand2 binding sites coincide with chromatin peaks specifically opened when Hand2 functions with GMT. Why has a potential role for Hand2 in pacemaker specification remained obscure? Since Hand2 knockout mice are embryonic lethal(Srivastava et al., 1997), pacemaker phenotypes would have been difficult to detect. However, re-examination of the Hand2 knockout phenotype should resolve this question, and the tools now exist to delete Hand2 in pacemaker cells. We speculate that Hand2 functions to establish the competence of pacemaker progenitor cells to undergo subsequent lineage specification.

Cardiac reprogramming provides a unique opportunity to perform structure-function analysis and to uncover important mechanistic insights, some of which may also apply to development. We found that the Hand2 N-terminus harbors a protein domain that ensures cardiac subtype diversification during cardiac reprogramming by GHMT. We envision two general models for how the CSD functions. First, the CSD regulates the binding specificity of Hand2 to promote cardiac subtype-specific gene expression signatures. Second, the CSD mediates key protein-protein interactions that are required for cardiac subtype diversity. To address the first possibility, we performed ATAC-seq

using Hand2Δ1-32 with GMT and found no significant differences in chromatin accessibility between full-length and mutant Hand2 (**Figure S5.8**). Thus, we currently favor the second model. We also found that Hand2 coordinates desmosome gene activation in iPMs. Interestingly, desmosomes are important for coordinating the assembly and function of gap junction proteins, which are essential for intercellular electrical communication(Corrado et al., 2017). Whether the insights gleaned from our studies have developmental importance remains an important question that now appears testable in parallel experimental systems for developmental and reprogramming cell-type specification.

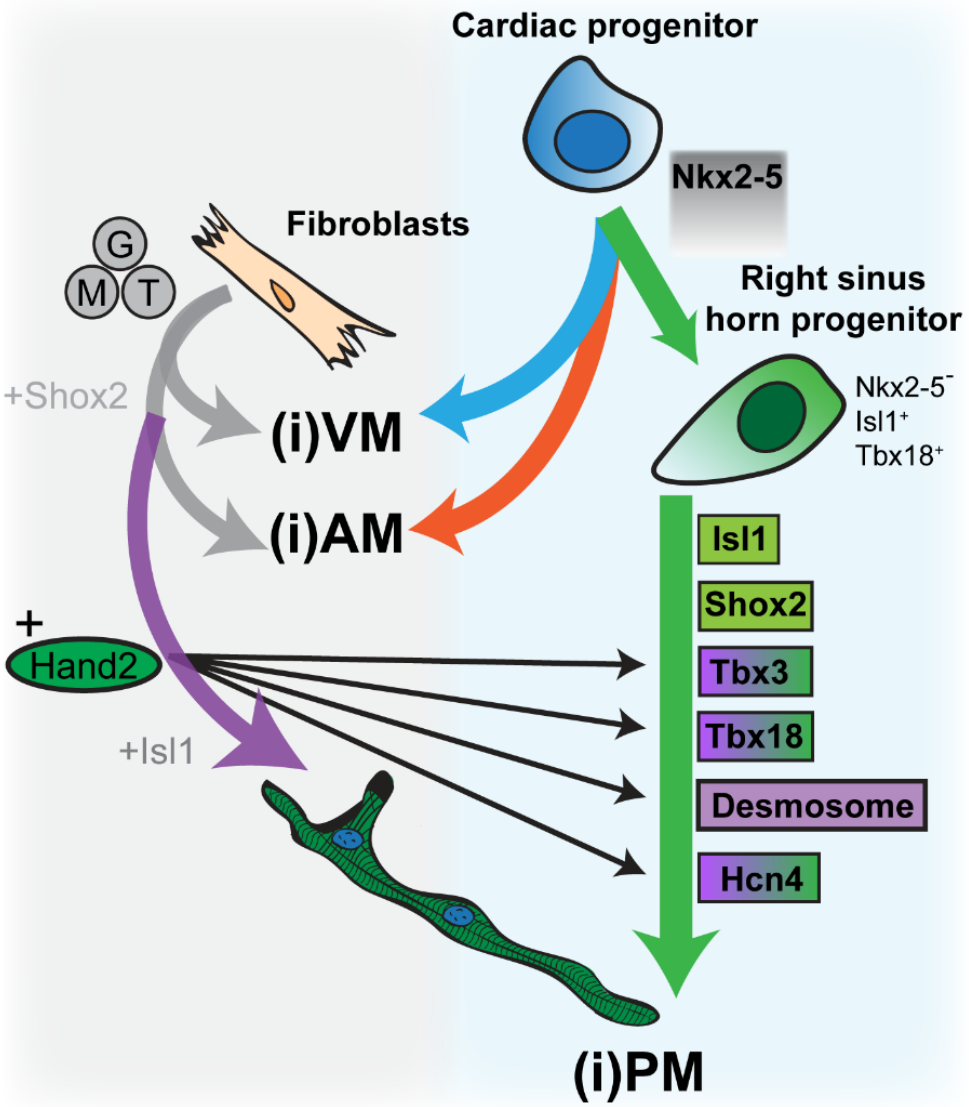


Figure 5.7. Model for Hand2-dependent mechanisms that regulate iPM formation by GHMT. *On the right, the pathway from cardiac progenitor cell to Tbx18⁺/Isl1⁺/Nkx2-5⁻ right sinus horn progenitor to endogenous pacemaker cell is shown. Previous work has established that Isl1, Shox2, Tbx3, and Tbx18 play instructive roles during developmental pacemaker formation. Cardiac progenitor cells can alternatively give rise to atrial myocytes (AM) and ventricular myocytes (VM) as depicted by orange and blue arrows, respectively. On the left is a proposed pathway for iPM reprogramming by GHMT. GMT alone is capable of converting fibroblasts into iAMs and iVMs (gray arrows), while GHMT augments iAM and iVM production and uniquely promotes iPM formation (purple arrow). In this study, we show that Hand2 partially activates the developmental pacemaker gene regulatory network via Tbx3 and Tbx8 (purple/green). However, Isl1 and Shox2 remain inactive following GHMT infection of fibroblasts (green), and exogenous Isl1 and Shox2 can promote iPM formation distinctly. Hand2 also coordinates activation of various core pacemaker marker genes, such as Hcn4 (purple/green), and we uncovered a novel iPM desmosome gene battery (purple).*

CHAPTER 5 ACKNOWLEDGEMENTS

We thank the UT Southwestern Medical Center Flow Cytometry, Histology Core, Shared Molecular Biology Confocal Microscopy, and Next Generation Sequencing Facilities for use of resources and expert technical assistance. We gratefully acknowledge Gary Hon, Ning Liu, and Beverly Rothermel for critical review of the manuscript and members of the Olson, Hon, and Munshi labs for technical advice, reagents, and scientific discussions. This work was supported by the NSF (Grant No.2015165336 to A.F.-P.), NIH (HL136604, HL133642, and HL135217 to N.V.M.), the Burroughs Wellcome Fund (1009838 to N.V.M.), and the March of Dimes Foundation (#5-FY13-203 to N.V.M.).

AUTHOR CONTRIBUTIONS

Conceptualization: N.V.M.; Experimental investigation: A.F.-P., M.B., and K.L.; Bioinformatic analysis: A.A.S. and C.X.; Writing -- Original Draft: N.V.M. and A.F.-P.; Writing -- Review & Editing: N.V.M. and A.F.-P.; Supervision: N.V.M. and C.X.; Funding acquisition: N.V.M. and A.F.-P.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

All animal experiments described in this study were conducted under the oversight of the University of Texas Southwestern Institutional Animal Care and Use Committee. Hcn4-GFP reporter mice were generated by bacterial artificial chromosome (BAC) transgenesis using a ~237 kb BAC clone (RP23-281H22) that includes the Hcn4 gene and nearby regulatory regions as previously described (Gong et al., 2003).

METHOD DETAILS

Isolation of mouse embryonic fibroblasts (MEFs)

Hcn4-GFP E12.5 reporter mouse embryos were harvested and removed of the head and internal organs. The remaining tissue was finely minced, and digested in 1 mL of 0.25% trypsin-EDTA per embryo for 15 min in a 37°C water bath. Cells were resuspended in 2 mL per embryo of fibroblast medium (DMEM/Hi glucose supplemented with 10% FBS and 1% [v/v] penicillin/streptomycin) and plated on a 15-cm dish. In 24 h, cells were harvested and stored for future use. All animal experiments described in this study were conducted under the oversight of the University of Texas Southwestern Institutional Animal Care and Use Committee.

DNA Constructs

The GFP-Hand2 fusion expression constructs were created by Phusion High-Fidelity PCR amplification of the appropriate fragment using primers containing BamHI, ClaI, and SalI linkers. Digested fragments were cloned into pBabe vector. For Hand2 domains, and N-Terminus variants, the desired fragments were amplified using primers that contained the desired truncation size. All constructs were confirmed via sequencing, expression via western blotting, and localization via immunofluorescence.

Generation of retroviruses, viral infection, and cardiac reprogramming

Two micrograms of retroviral plasmid DNA mix expressing mouse Gata4, Mef2c, Tbx5, Hand2 , or Hand2 variants, were transfected using FuGENE 6 into Platinum E cells

plated on 6-well plate at a density of 1 x 10^6 cell per well 24h post-plating(Song et al., 2012c). In parallel, Hcn4-GFP MEFs were seeded into culture dishes or plates that were precoated with SureCoat (Cellutron) for 2 h at a density of 6 x 10^4 per square centimeter. 24 h post-transfection, viral supernatant was collected and filtered through a 0.45 μm cellulose filter. The viral medium was mixed with hexadimethrine bromide (polybrene) at a final concentration of 8 μg/mL and added to the culture plate with MEFs. Platinum E cells were replenished with growth medium. 24 h later, MEFs were infected with a second round of viral supernatant, and Platinum E cells were properly discarded. 48 h post-infection, the medium was replaced with induction medium composed of DMEM/199 (4:1), 10% FBS, 5% HI Horse serum. 1% penicillin/streptomycin, 1% non-essential amino acids, 1% essential amino acids, 1% B27, 1% insulin-selenium-transferrin, 1% vitamin mixture, 1% sodium pyruvate (Invitrogen). The medium was replaced every 2-3 days until the end of the experiment.

Immunocytochemistry

Cells were washed two times with PBS and fixed with 4% paraformaldehyde for 15 min and permeabilized with 0.1% Triton X-100 at room temperature. Samples were blocked with universal blocking buffer (BioGenex) for 10 mins at RT and then incubated with primary antibodies against α-actinin (Sigma), GFP (Invitrogen), Nppa (Abgent), or Myl2 (ProteinTech) diluted in 0.5X universal blocking buffer overnight at 4ºC. The next day, samples were washed three times with 0.1% Triton X-100, and incubated with goat anti-mouse Alexa Fluor 555 (Invitrogen), goat anti-chicken Alexa Fluor 488 (Invitrogen), and donkey anti-rabbit Alexa Fluor 647 (Invitrogen) for 1 h at room temperature. Cells

were washed three times with 0.1% Triton X-100, and mounted with ProLong Glass antifade mountant with NucBlue (Invitrogen). Images where capture with either a Zeiss LSM700 or Nikon A1R+ confocal microscope.

We previously identified a set of minimal immunostaining markers that distinguish endogenous atrial, ventricular, and pacemaker myocytes. This analysis demonstrated that endogenous pacemaker cells are sarcomere^{+}/Hcn4-GFP^{+}/Nppa^{-}/Myl2^{-}. We have also shown that this combination of immunostaining and morphological criteria prospectively identifies reprogrammed iPMs as confirmed by patch clamp recordings(Nam et al., 2014). Thus, we feel confident that our parsimonious choice of markers accurately identifies functional pacemaker cells

Luciferase assays

Expression constructs (pcDNA3.1-mycHand2, pcDNA3.1-mycHand2, pcDNA3.1-mycHand2-Variants) and reporters constructs (α-MHC and ANF-Luciferase) were co-transfected into PE cells using Fugene6 (Promega) as per manufacturer's instructions. DNA concentration was kept constant by filling with empty expression vector. To normalize expression, samples were also co-transfected with CMV-LacZ plasmid. The luciferase activity was measured 48 h post-transfection using the luciferase assay kit (Promega) according to manufacturer's instructions. β–galactosidase activity was used to calculate relative expression. All experiments were carried out in triplicate and performed at least three separate times.

Real-time PCR

Total RNA was extracted using RNeasy Mini kit (Qiagen) according to manufacturer's instructions. One microgram of RNA was reverse-transcribed using SuperScript III (Invitrogen). AzuraQuant (Azura Genomics) RT-qPCR was performed in triplicate using an ABI PRISM real-time PCR machine (Applied Biosystems) and analyzed with SDS2.45 Software (Applied Biosystems). mRNA levels were normalized to Gapdh.

Western blotting and co-immunoprecipitation

Platinum-E cells were transiently transfected with myc-epitope or GFP fusion Gata4, Mef2c, Tbx5, Hand2, or Hand2 variants. 24 and 48 h post-transfection, 10T1/2 cells were infected with viral supernatant as previously described. 48 h post-infection 10T1/2 were collected and protein lysate prepared using three freeze-thaw cycles in RIPA buffer (50 mM HEPES-KOH, Ph 7.5, 1 mM EDTA, pH 8.0, 1 %[v/v] NP-40 Igepal CA-630, 0.7% Sodium deoxycholate, 0.5M Lithium Chloride, 1X protein inhibitor). Cell debris was cleared by centrifugation at 2000x *g* for 5 min at 4ºC. Supernatant was collected and used for downstream applications. Twenty micrograms of total protein lysate (as determined by Bradford assay) were resolved by SDS/PAGE, transferred to PDVF membrane (BioRad), probe with mouse anti-myc (Thermo Fisher Scientific), or rabbit anti-GFP (Invitrogen), and detected with HRP-conjugated anti-mouse/rabbit. Membranes were incubated with Clarity MAX western ECL blotting substrates (BioRAD) according to manufactures instruction and developed using Licor's Odyssey Fc imaging system. For Co-IP experiments, PE cells were co-transfected with GFP-Tbx5 and myc-Hand2 variants

for 48hrs. Cell lysates were incubated overnight with 1 µg of rabbit polyclonal anti-GFP (Thermo Fisher Scientific) with end-over-end rotation at 4ºC. The cell lysates were pulled down using magnetic Protein G Dynabeads (Invitrogen), and then the myc-Hand2 was eluted with 100 mM Glycine, pH 2, for 10 min at 55ºC. Solution was neutralized with 10µL of 1M Tris, pH 8.0. Eluted and input samples were analyzed by SDS-PAGE Western blot using a GFP polyclonal antibody or a mouse monoclonal anti-Myc antibody.

Flow cytometry

For isolation of GFP-Hand2^{+}Hcn4-GFP^{+} populations after 14 days of reprogramming, cells were trypsinized and washed twice with FACS buffer (PBS, 0.5%BSA, 0.1% NaN_3), resuspended at 1 x 10^6 cells per mL in FACS buffer and filtered with a 100 µm cell strainer. Cells were sorted using a FACSAria 2 (BD Biosciences). For ATAC-seq samples, 12.5k GFP-Hand2^{+}Hcn4-GFP^{+} cells were isolated per biological replicates, and ~40k cells for RNA samples. Separate plates were considered as biological samples and were not pooled.

RNA-seq library preparation and sequencing

RNA-seq was performed in biological duplicates. For RNA-seq, total RNA was extracted from sorted GFP-Hand2^{+}Hcn4-GFP^{+} cells using the RNeasy Mini kit (Qiagen). 100 ng of total RNA was used for the RNA-seq. Library preparation was performed using TruSeq Stranded mRNA with Ribo-Zero human/mouse/rat assay (Illumina) following the manufacture's protocol. Optimal library size was confirmed using Agilent 2200

TapeStation using D1000 screen tape. Samples were then paired-end sequenced using an Illumina HiSeq 500 High Output (400M) at the University of Texas Southwestern Microarray Core Facility.

RNA-seq analysis

Raw FASTQ files were analyzed using FastQC v0.11.2(Andrews, 2010) and FastQ Screen v0.4.4(Wingett, 2011), and reads were quality-trimmed using fastq-mcf (ea-utils/1.1.2-806)(Aronesty, 2013). The trimmed reads were mapped to the hg19 assembly of the human genome (the University of California, Santa Cruz, version from igenomes) using Tophat(Kim et al., 2013). Duplicated reads were marked using Picard tools (v1.127; https://broadinstitute.github.io/picard/), the RNA counts generated from FeatureCounts(Liao et al., 2013) were TMM normalized(Robinson and Oshlack, 2010), and differential expression analysis was performed using edgeR(Robinson et al., 2010). For differential expression analysis, statistical cutoffs of FDR ≤ 0.05 and log2CPM ≥ 0 and were used to identify statistically significant and possibly biologically relevant differentially regulated genes. Mouse sinoatrial node transcriptome raw data was downloaded from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65658 (P4, E14, E14.5 for SAN and RA samples) and counting and normalization was done as above. Genes common to each cluster from the RNASEQ heatmap and also significantly differentially expressed between SAN versus RA for E14.5, E14 and P4 datasets were plotted in venn charts generated using Intervene(Khan and Mathelier, 2017).

Pathway and network analysis were conducted using the Database for Annotation, Visualization, and Integrated Discovery (DAVID v.6.8)(Huang et al., 2008) software. The threshold was set as modified Fisher Exact P-value (EASE score) ≤ 0.05. Differentially expressed gene heat maps were clustered by hierarchical clustering using R (http://www.R-project.org) and generated using the heatmap.2 function within the gplots package(Warnes et al., 2016) and ComplexHeatmap package(Gu et al., 2016).

Gene ontology terms were generated using the web-based PANTHER Gene Ontology (GO) tool using nearest neighbor gene for common peaks within replicates. GO clusters and mini RNA-seq heatmaps were generated using the ggplot2 R package (http://ggplot2.org/).

ATAC-seq library preparation and sequencing

ATAC-seq was performed in biological duplicates. To profile low amounts of open chromatin in our samples we used a modified ATAC-seq protocol previously published (Corces et al., 2017). In brief, 12.5k GFP-Hand2^{+}Hcn4-GFP^{+} cells were sorted and pelleted by centrifugation at 500 x *g* at 4ºC for 10 min. Cells were resuspended in 50 µL of cold RBS (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$ in H_20) with 0.1%NP40, 0.1%Tween-20, and 0.01% Digitonin, and incubated on ice for 3 min. The isolated nuclei was washed one time with 1mL of cold RSB containing 0.1% Tween-20 but not NP40 or Digitonin. Nuclei were pelleted by centrifugation at 500x *g* at 4ºC for 10 min before the supernatant was carefully discarded. The isolated nuclei was resuspended in 50 µL of transposition mixture (1xTD Buffer, 100nM Tn5, PBS, 0.01% Digitonin, 0.1% Tween-20,

H20) and incubated at 37ºC for 30 min in a thermomixer at 1000 RPM mixing. The DNA was cleaned with Zymo DNA clean and concentrator-5 (Zymo Research), and eluted in 20 µL of Elution buffer. Illumina adaptor Indexes were added to the libraries using KAPA HiFi enzyme with manufacture's PCR conditions (Kapa Biosystems) and pre-amplified for 5 cycles. One µL of this reaction was run on the Kapa library quantification kit (Kapa Biosystems) to estimate leftover cycles for a final concentration of 20 nM per library. Amplified DNA was cleanup twice using AMPure XP Beads (Beckman Coulter); the first using a 0.5X volume to remove large DNA fragments (>1kb), and the second with 1.2X to clean up the remaining small fragments. The final clean fragments were resuspended in 20 µL of Elution buffer. DNA was quantify using a Qubit fluorometer (life technologies), and library sizes were determined with a TapeStation (Agilent Technologies). Sequencing was performed using the Illumina HiSeq 500 High Output (400M) to obtain an average of 50 million paired-end 75bp reads per sample.

ATAC-seq analysis

Raw FASTQ files were analyzed using FastQC v0.11.5(Andrews, 2010) and FastQ Screen v0.11.4(Wingett, 2011), and reads were quality-trimmed using fastq-mcf (ea-utils/1.1.2-806)(Aronesty, 2013). The trimmed reads were mapped to the mm10 assembly of the mouse genome (the University of California, Santa Cruz, version from igenomes) with bowtie2 (version 2.3.3.1)(Langmead and Salzberg, 2012). The duplicates were marked using picard-tools (v2.2.1) and blacklist regions were removed using bedtools (v2.7.1). TN5 shifting of bam files was performed using the open-source Perl script "ATAC_BAM_shiftrt_gappedAlign.pl"(Ahmed and Ucar, 2017). The ATAC-Seq peaks

were called using MACS2 (version 2.1.0.20160309)(Zhang et al., 2008), with a q-value threshold of 0.05 and using the random background. Differential binding analysis was performed using the DiffBind package(Stark and Brown, 2011).

Furthermore, the fragments were divided as nucleosome free (sub-nucleosomal, <100 bp) and as nucleosome occupied or bound (>100 bp). Overlapping peaks from replicates were merged using bedtools (v2.7.1) merge to yield a peak set for each sample. We used bedtools (v2.7.1) intersect to get common regions between peak files of replicates of the same experimental group and subsequently filtered for all peaks which were consistently and significantly different (FDR < 0.05) between at least two conditions. The annotation of peaks, genome-wide distribution of ATAC-Seq regions on promoters, exons, introns, and intergenic regions in each of the samples was done using ChIPseeker(Yu et al., 2015).

For ATAC-seq in model organisms, the peak file (NAME_peaks.narrowPeak) can be uploaded directly to the UCSC genome browser for generating the browser tracks and ENCODE peaks were added by configuring the UCSC Genome Browser Track to import public track hubs, in this case ENCODE ChIP-Seq dataset from mouse limb https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE55707 (GSM1447340 – (ChIPseq), GSM1447341- (Input)) and also mouse heart https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE73368. The mm9 genome based wig files were converted to mm10 genome bigwig files using CrossMap. For genome browser tracks bigWig files were hosted in Cyverse(Merchant et al., 2016) Web-

based PANTHER Gene Ontology (GO) tool was used for generating GO terms for nearest neighbor genes for peaks that were common to 2 replicates of each ATAC-seq sample. We generated Pearson correlation graphs as well as RPKM normalized bigwig files (bamCoverage module), PCA plots, heatmaps and profile plots for the ATAC-Seq using deepTools2(Ramirez et al., 2016). For ATAC-seq profile plots, ATAC-seq counts were first normalized by RPKM and then z-scores were calculated on a per-peak basis across samples. RPKM in this case is Reads per Kilobase per Million mapped reads.

The method mentioned in Laurent et. al.(Laurent et al., 2017a) for motif analysis was followed in the ATACSeq dataset. De-novo and known motif over-representation analysis in the selected merged peaksets was done using the findMotifsGenome.pl script in the HOMER collection(Heinz et al., 2010). A fixed length of 7 nucleotides was used for de-novo motif finding in +/- 150 bp region around the summit of each peak. We used FIMO(Grant et al., 2011) from the MEME suite to confirm the discovery of de-novo motifs. The top 5 over-represented motifs with all possible alternative sequences were used to scan the region for high-affinity matches with P-value threshold of 1e-4.

ChIP-seq and ATAC-seq integration analysis

1684 regions from merged ATAC-seq data (for the figure (ATACSEQ final heatmap) that overlapped with ChIP-seq Peaks from Laurent et al.(Laurent et al., 2017a) were found using bedtools (v2.7.1) intersect. De-novo motif finding was performed in these regions as above and FIMO was used to scan motifs for GATA factors 1, 2, 3, 4 and 6; TBX factors 2, 3 and 5 as well as MEF2 factors A, C and D. Hierarchically clustered

heatmaps representing the patterns of these TF motifs as also all possible E-box motif (CANNTG) combinations against the ChIP-seq and ATAC-seq integrated regions was plotted using the ComplexHeatmap package(Gu et al., 2016).

RNA-seq and ATAC-seq integration analysis

Fold changes (log to the base 2) of significant differentially expressed genes (FDR < 0.05) were calculated for RNASEQ and corresponding ATACSEQ peaks annotated closest to the TSS of these genes by ChIPseeker for GHMT versus GMT conditions. A scatterplot was generated using this dataset with base R (http://www.R-project.org).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

All data are represented as mean ± S.D. and have at least $n = 3$ per group (refer to figure legend to detailed information). P values were calculated with either unpaired *t*-test or one-way ANOVA except where noted in the Figure Legend. Statistical analysis was run using GraphPad Prism 7.04 software package. $P < 0.05$ was considered significant in all cases. P values are depicted using the GP style with one to four asterisks (<0.05, <0.01, <0.001, <0.0001).

Data and Software Availability

The accession number for the ATAC-seq, and RNA-seq data in this paper is GEO: GSE124338.

SUPPLEMENTAL FIGURE LEGENDS

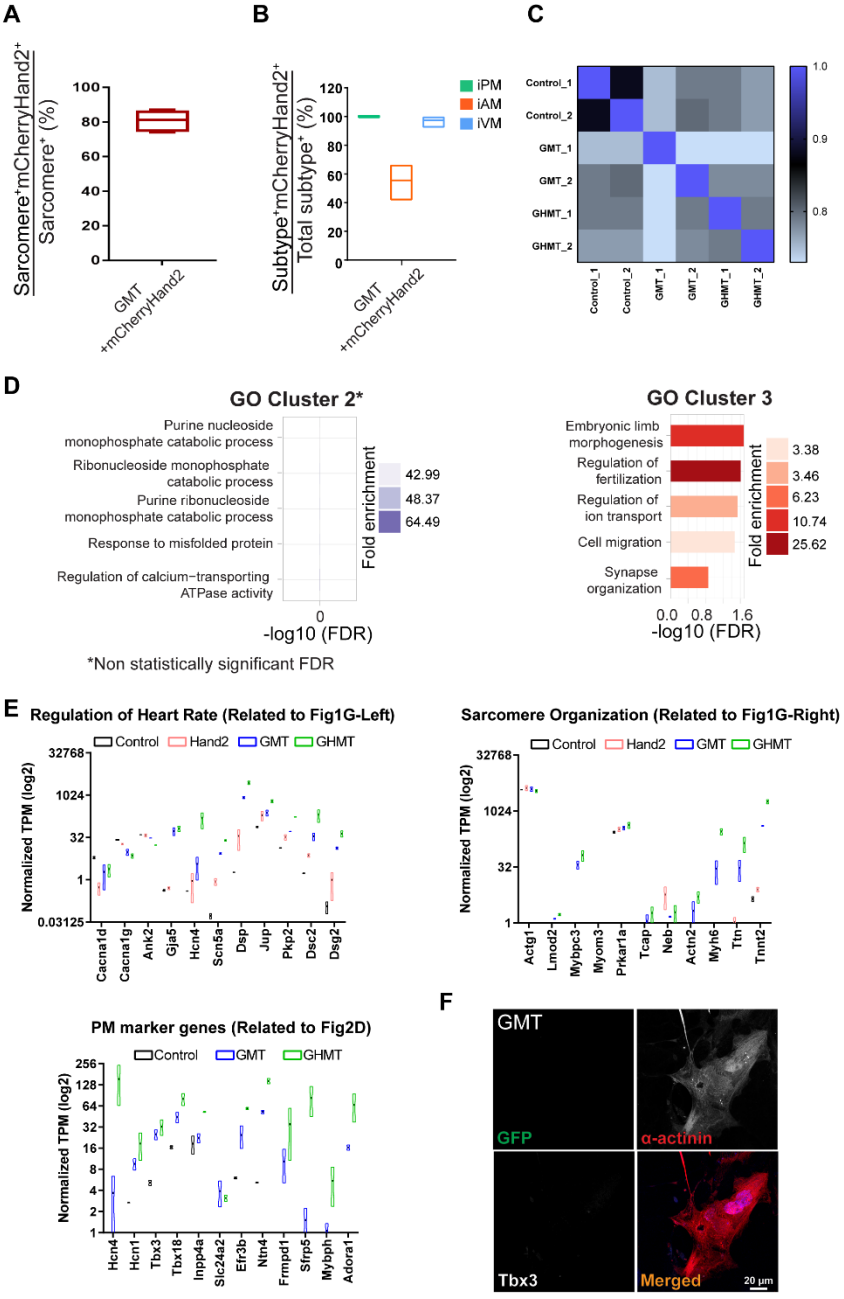


Figure S1. Additional cardiac reprogramming quantification and RNA-seq quality control analysis. *A) Hcn4-GFP MEFs were infected with GMT and a Hand2-mCherry fusion construct to quantify iCLM formation specifically in Hand2⁺ cells (mean±SD, n=3). B) Same as in (A) for iPMs, iAMs, and iVMs (mean±SD, n=3). C) Correlation matrix for all RNA-seq samples including biological duplicates. Pearson correlation is indicated by color legend at right. Raw values are provided in Table S1. D) Gene ontology (GO) term analysis for clusters 2 and 3 from Figure 1D. E) Violin plots of RNA-seq data corresponding to Figures 1G and 2D. F) Immunocytochemistry (ICC) on GMT-reprogrammed cells for Tbx3. Scale bar, 20µm.*

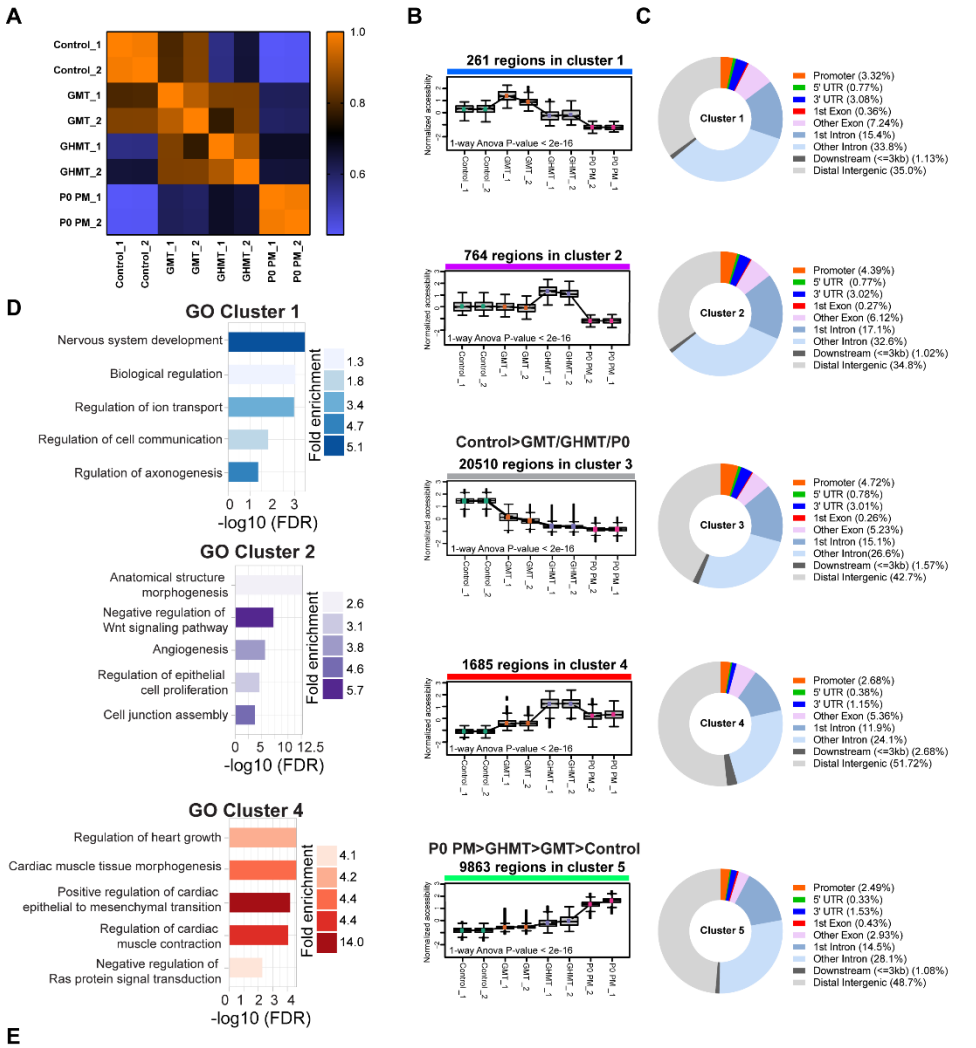


Figure S2. Additional ATAC-seq data analysis. *A) Correlation matrix for all ATAC-seq samples including biological duplicates. Pearson correlation is indicated by color legend at right. Raw values are provided in Table S3. B) Aggregate chromatin accessibility for each ATAC-seq sample in clusters 1-5 from Figure 3A. p-values for 1-way ANOVA testing across samples are shown within individual box plots. Statistical testing for pairwise comparisons within a given cluster was performed by Tukey's HSD test between groups. All comparisons were significant, and numerical values for Tukey's HSD testing are provided in Table S2. C) Genomic location of the peaks in clusters 1-5 from Figure 3A. D) Gene ontology (GO) term analysis for clusters 1, 2, and 4 shown in Figure 3A. E) HOMER analysis to identify binding motifs for clusters 1, 2, and 4 from Figure 3A.*

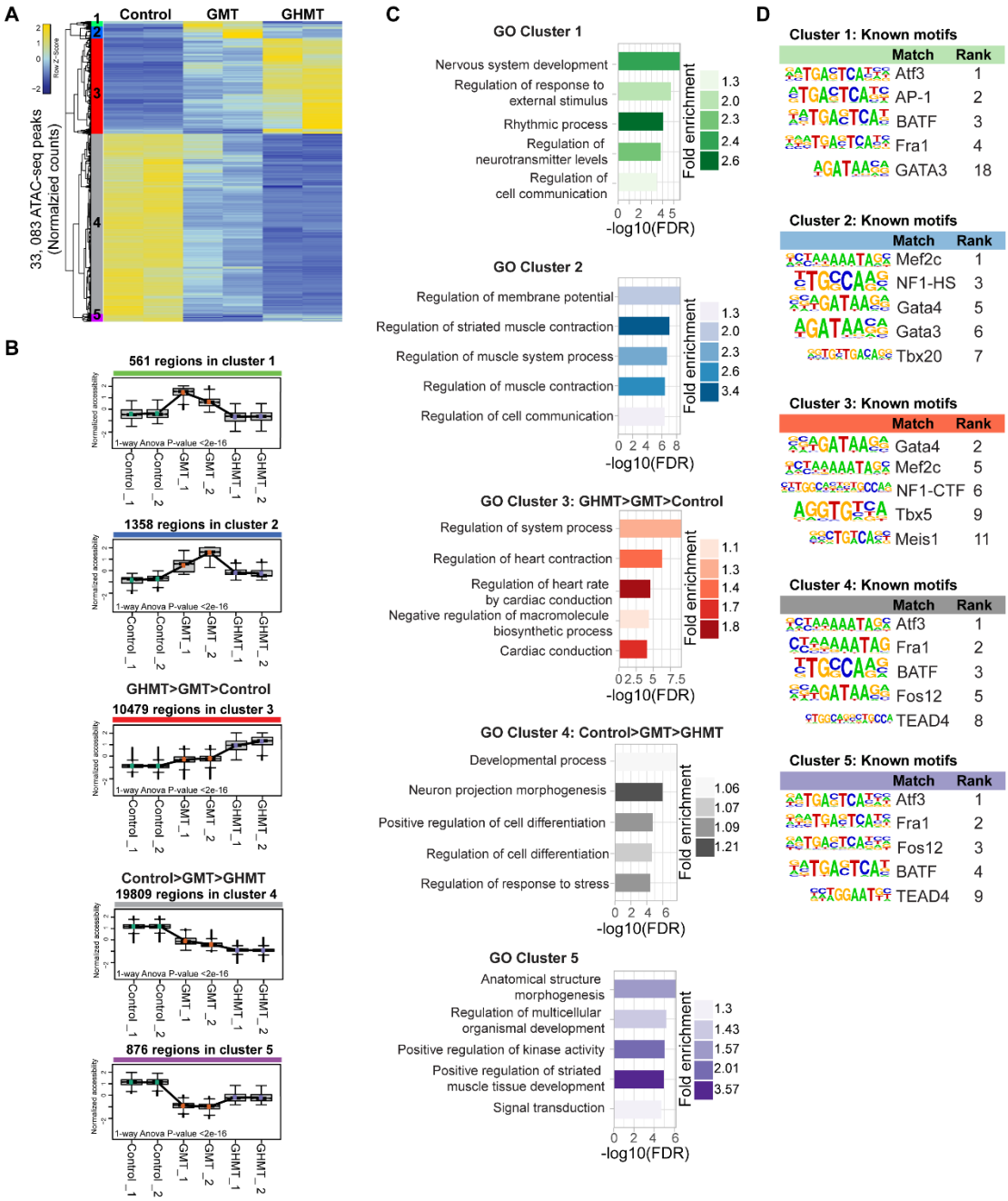


Figure S3. Analysis of ATAC-seq data after excluding endogenous P0 PMs. *A) Heatmap for the three-way comparison of ATAC-seq datasets derived from Control, GMT, and GHMT samples after unsupervised clustering (Clusters 1-5). The peaks used in this analysis are the same as those that comprise the heatmap shown in Figure 3A. B) Aggregate chromatin accessibility for each ATAC-seq sample in clusters 1-5 from (A). p-values for 1-way ANOVA testing across samples are shown within individual box plots. Statistical testing for pairwise comparisons within a given cluster was performed by Tukey's HSD test between groups. All comparisons were significant, and numerical values for Tukey's HSD testing are provided in Table S2. C) Gene ontology (GO) term analysis for clusters shown in (A). D) HOMER analysis to identify binding motifs for clusters shown in (A).*

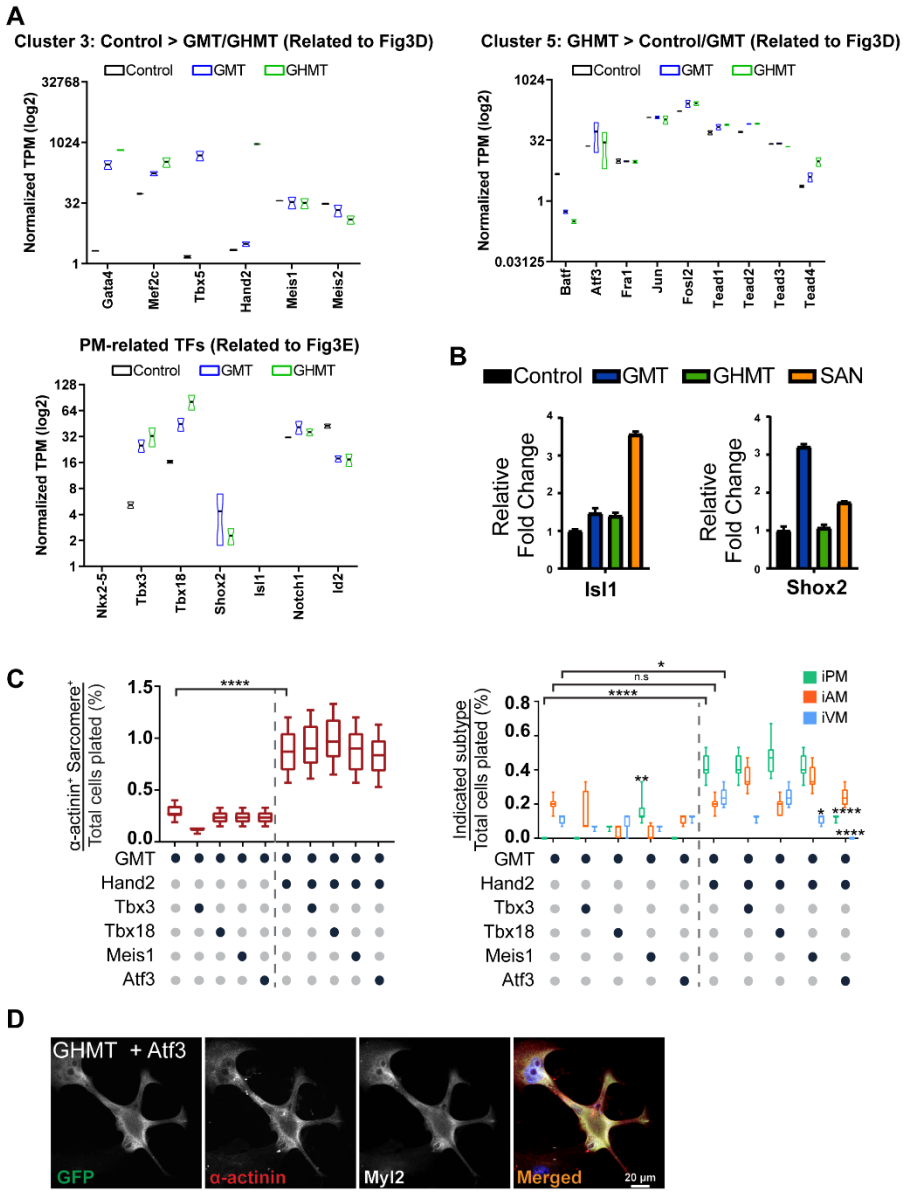


Figure S4. Inclusion of specific factors and their effect on cardiac reprogramming. *A) Violin plots of RNA-seq data corresponding to Figures 3D and 3E. B) qRT-PCR analysis of reprogrammed cells confirms that expression of Isl1 and Shox2 are not activated in iPMs. C) Specific transcription factor candidates identified in Figure 3C-D were tested by cardiac reprogramming assay for sarcomere organization (left) and subtype diversification (right). All samples were compared with GMT (left of dashed line) or GHMT (right of dashed line) to determine statistical significance. Additional pairwise comparisons are shown with horizontal lines at the top of each graph. D) Immunocytochemistry (ICC) on GMT-reprogrammed cells for Tbx3. Scale bar, 20μm. D) Representative example of a mixed Hcn4-GFP⁺/Myl2⁺ iCLM observed during cardiac reprogramming with GHMT + Atf3. Statistical significance indicated as follows: *, p<0.05; **, p<0.01; ***, p<0.001; ****, p<0.0001. Non-significant comparisons indicated either by "ns" or no asterisk.*

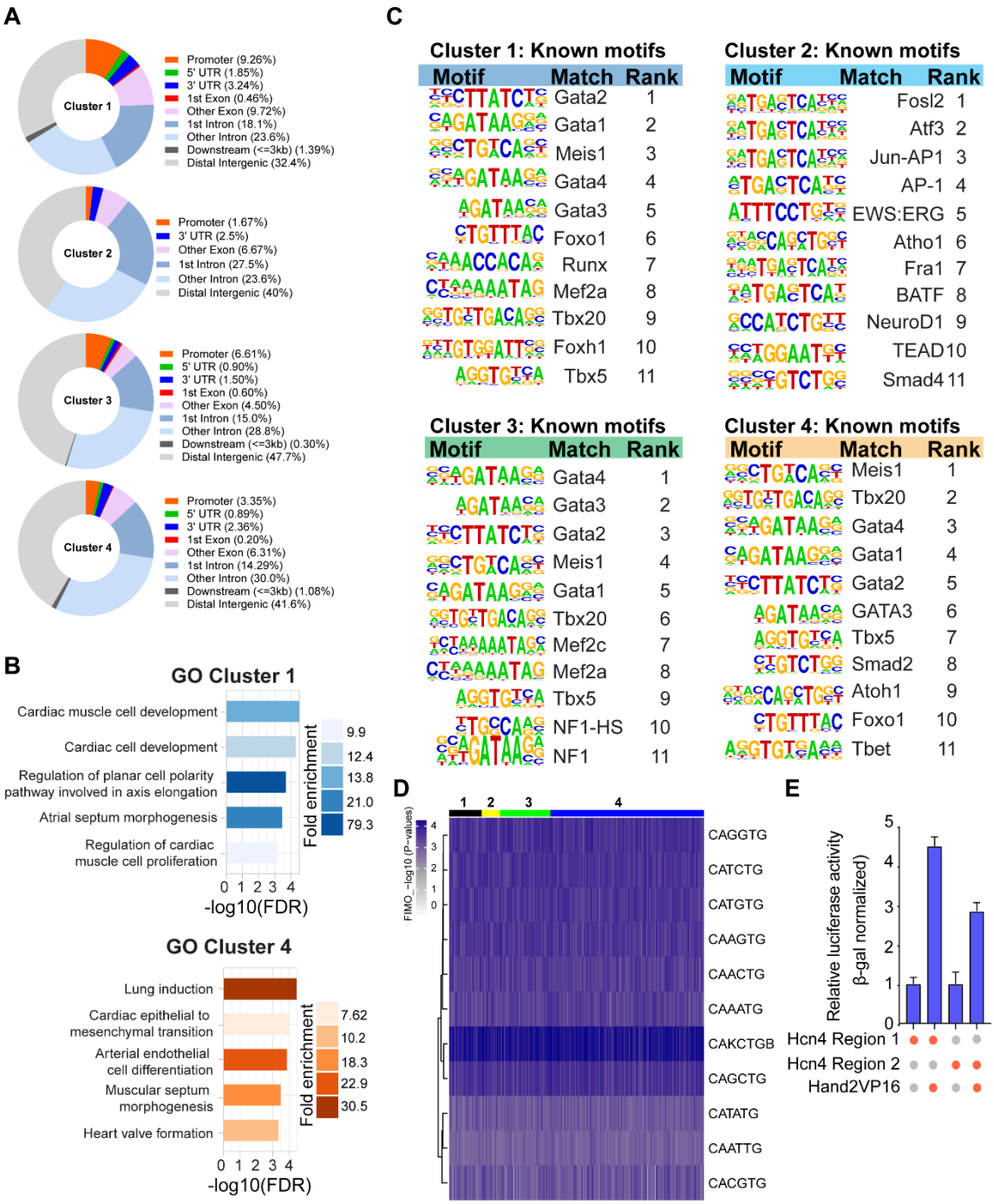


Figure S5. Additional analysis of the intersectional dataset derived from Hand2 ChIP-seq and GHMT ATAC-seq. *A) Genomic location of the peaks in clusters 1-4 from Figure 4A. B) Gene ontology (GO) term analysis for clusters 1 and 4 shown in Figure 4A. C) Known transcription factor binding motif identification for clusters 1-4 from Figure 4A. D) Distribution of E-box CANNTG biding motifs across the iPM Hand2 cistrome. E) Luciferase assay from transient transfection of HEK cells with the indicated Hcn4 region reporters with or without a Hand2-VP16 expression construct.*

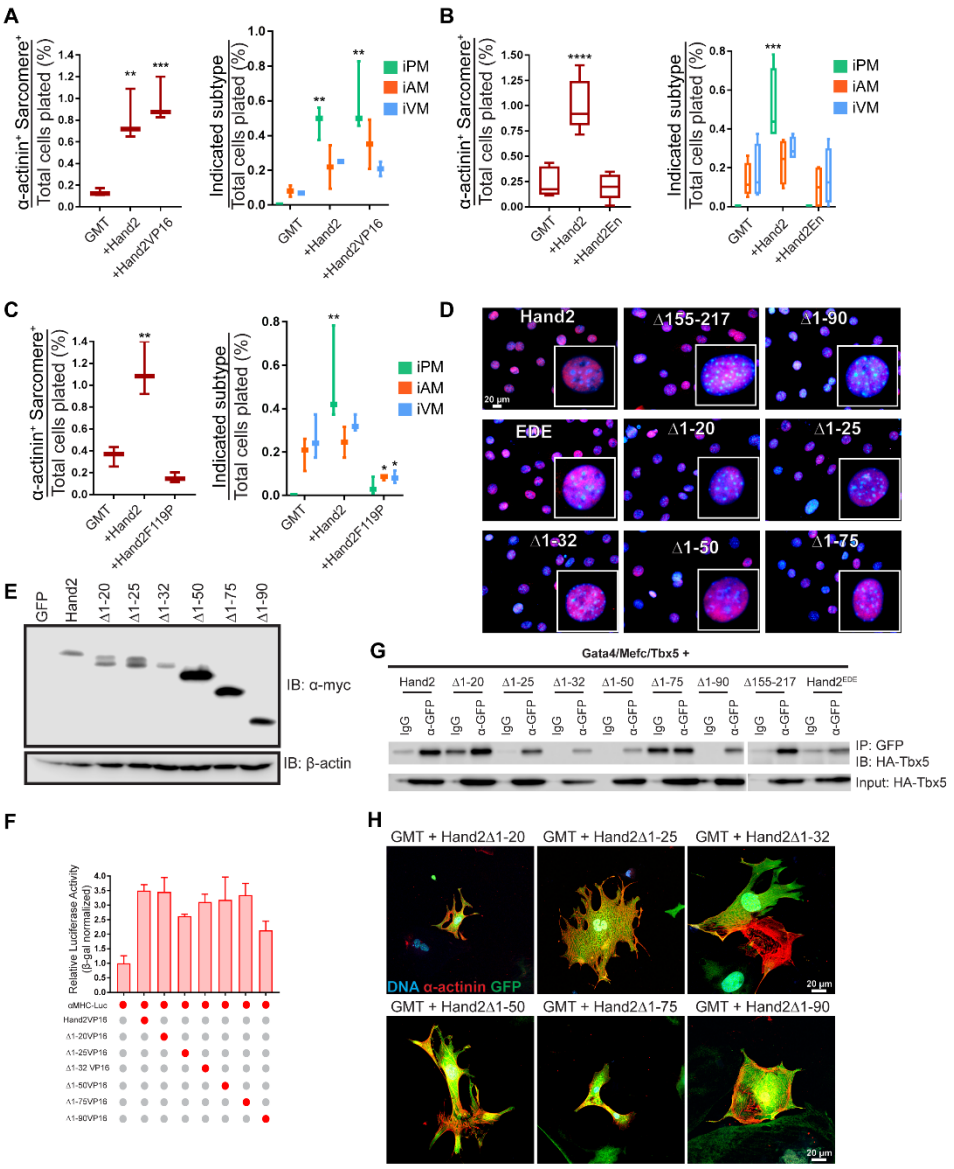


Figure S6. Hand2 function during cardiac reprogramming and fidelity of Hand2 mutants. *Cardiac reprogramming was performed in the presence of Hand2-VP16 (A), Hand2-EnR (B), and Hand2^{F119P} (C). Sarcomere organization (left) and subtype diversification (right) was assessed for each (mean±SD, n=3). All samples were compared with GMT to determine statistical significance. D) Nuclear localization of Hand2 mutants was confirmed by immunocytochemistry (ICC) and confocal imaging. Scale bar, 20□m. E) Production of full-length Hand2 mutants was validated by Western Blot. F) DNA binding capacity of Hand2 mutants was assessed by luciferase assay in the presence of □MHC-Luc reporter and VP16 fusion proteins. G) Protein-protein interaction capacity of Hand2 mutants was tested by co-immunoprecipitation analysis with Tbx5. H) ICC analysis of cardiac reprogramming by GMT plus the indicated Hand2 mutants. Scale bar, 20□m. Statistical significance indicated as follows: *, p<0.05; **, p<0.01; ***, p<0.001; ****, p<0.0001. Non-significant comparisons indicated either by "ns" or no asterisk.*

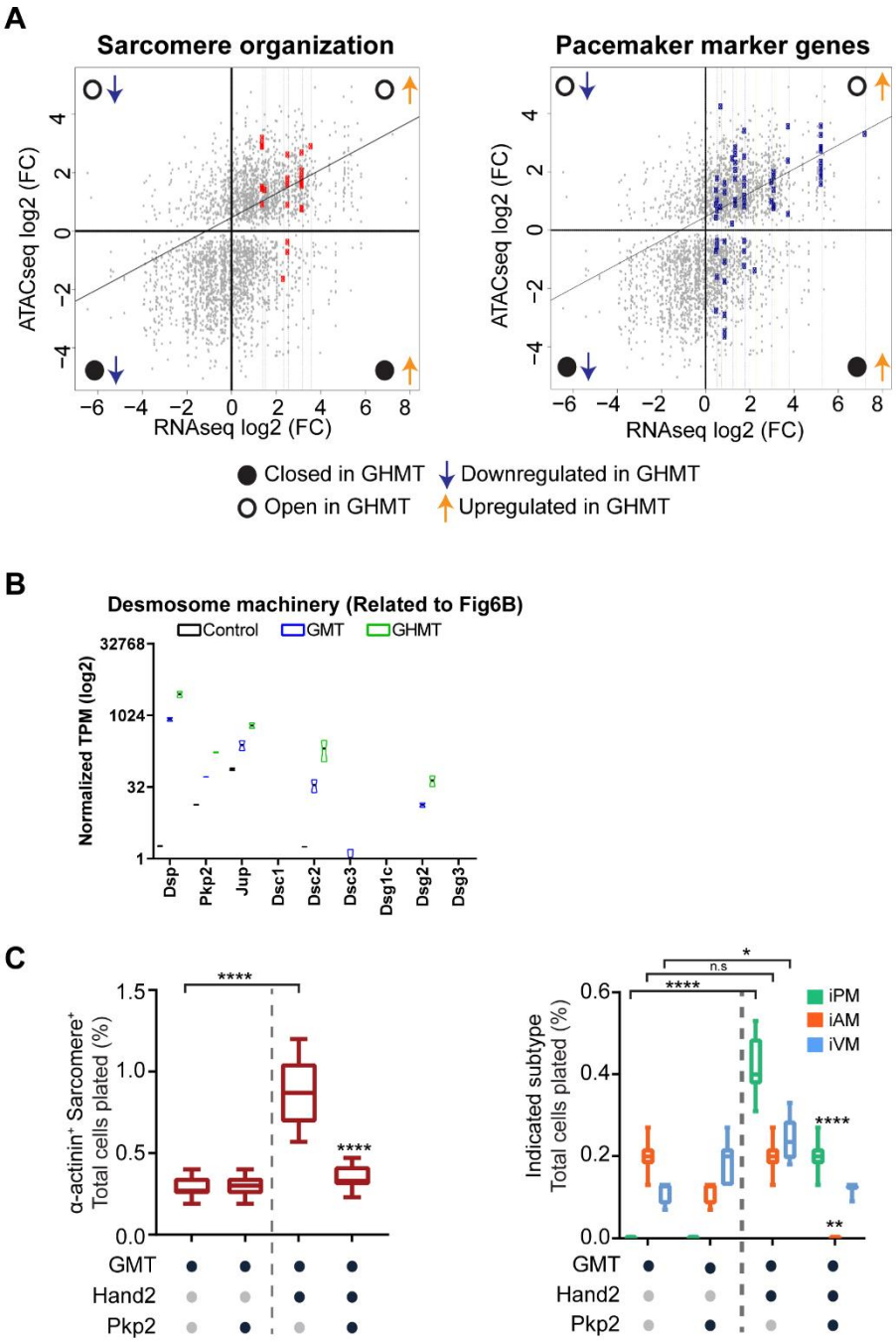


Figure S7. Additional analysis of integrated ATAC-seq and RNA-seq data. *A) Integrated analysis as in Figure 6A with sarcomere genes highlighted in red (left) and pacemaker marker genes highlighted in blue (right). B) Violin plots of RNA-seq data corresponding to Figure 6B. C) Pkp2 was tested for its activity in a cardiac reprogramming assay for sarcomere organization (left) and subtype diversification (right) (mean±SD, n=6). Samples were compared with GMT (left of dashed line) or GHMT (right of dashed line) to determine statistical significance. Additional pairwise comparisons are shown with horizontal lines at the top of each graph. Statistical significance indicated as follows: *, p<0.05; **, p<0.01; ***, p<0.001; ****, p<0.0001. Non-significant comparisons indicated either by "ns" or no asterisk.*

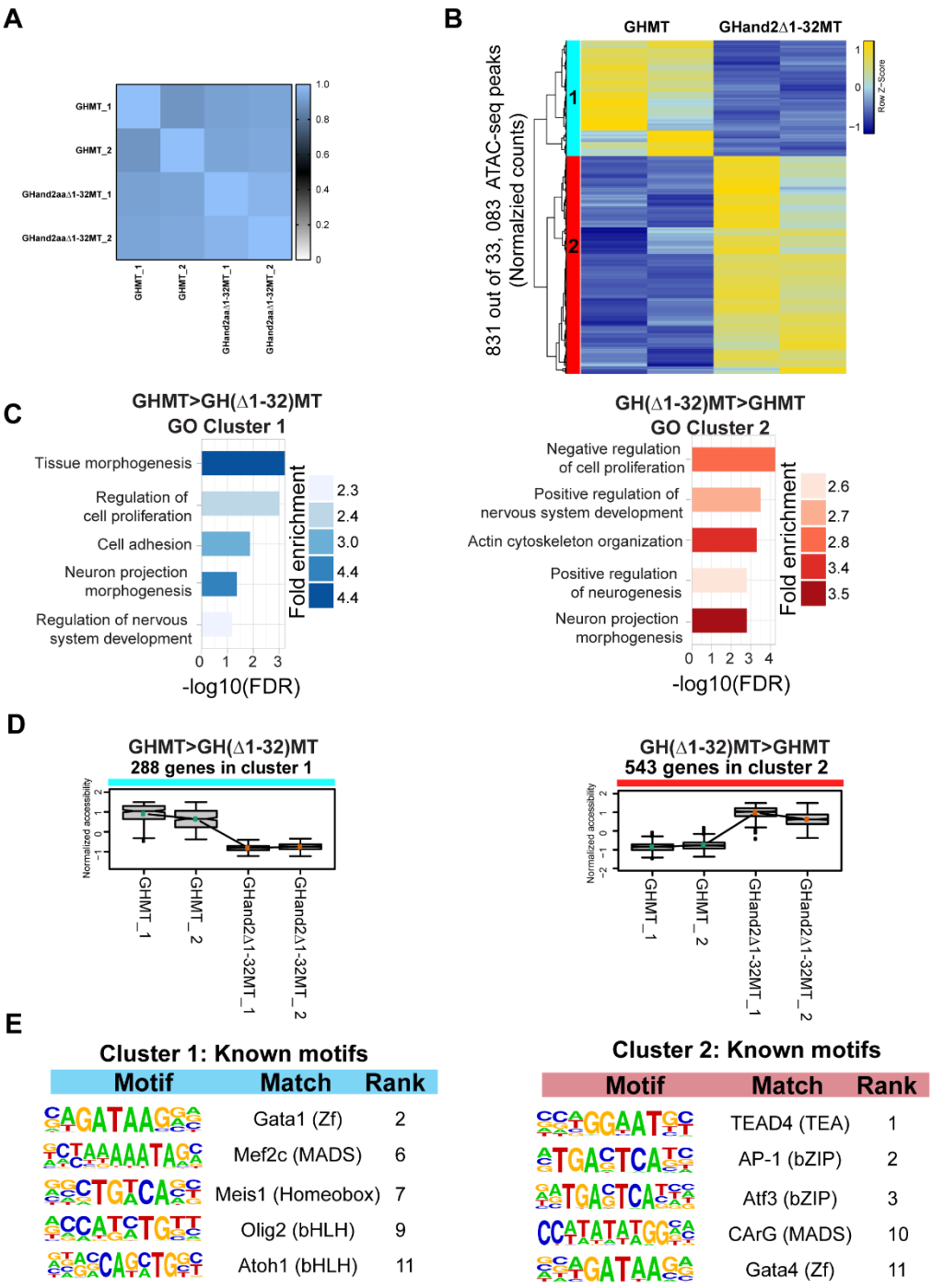


Figure S8. The Hand2 CSD domain does not significantly influence chromatin accessibility. *A) Correlation matrix for ATAC-seq libraries generated from GMT plus Hand2 and GMT plus Hand2Δ1-32 samples. Pearson correlation is indicated by color legend. Raw values are provided in Table S17. B) Heatmap showing the 831 out of 33,083 loci from Figure 3A that displayed differential chromatin accessibility between the GHMT and GMT plus Hand2Δ1-32 samples. C) Gene ontology (GO) term analysis for clusters 1 and 2. D) Aggregate chromatin accessibility for clusters 1 and 2. E) Predicted binding motifs for clusters 1 and 2.*

CHAPTER SIX

FINAL REMARKS AND FUTURE PERSPECTIVES

Chapter overview

In this final chapter I will briefly summarize the contributions of this thesis and discuss future directions that the next scientist could consider in moving forward with future cardiac reprogramming projects.

Impact of this doctoral thesis

In this thesis, I have optimized and applied a series of methods for the direct reprogramming of mouse embryonic fibroblasts into iCM. Using ATAC-seq, RNA-seq, transcription factor domain mutants, and immunofluorescent quantification I was able to 1) document the difference between GMT- and GHMT-chromatin and gene expression changes during iCM development of Hcn4-GFP^{+} populations, 2) dissect the Hand2-dependent functions in the context of GMT, and 3) Identify novel Hand2 domains responsible for the regulation of iPM cell fate formation. Focusing on GHMT-iPMs, we observe some shared transcriptional networks with endogenous sinoatrial tissues (Tbx3/Tbx18). Although, some canonical networks are still missing (Isl1/Shox2). Focusing on Hand2-dependent cistrome changes, we observed an increase accessibility that, in combination with GMT, are able to promote both structural cardiac genes, SAN specific gene expression, and a novel role of the desmosome machinery (regulator of cell-cell communication). We further dissect Hand2 into functional domains and identify a novel cardiac subtype diversity (CSD) domain within the N-terminus low complexity regions.

Together, this works applied a unique, stringent, and controlled model system that illuminated the mechanism by which Hand2 enabled iPM formation. This approach can pave advances in “off the shelf” PM cellular engineering.

Perspective on how to move forward

Currently, we utilize mouse embryonic fibroblasts as a source for iCM reprogramming. Although this provide us with a constant source of primary cells, and the capacity for interesting genetics, the translation impact of murine iCM-iPM formation towards human cell populations is still been studied. Presently, GHMT is unable to reprogram human dermal fibroblasts without the use of additional factors MYOCD, miR-1, and miR-133 (Nam et al., 2013). However, reprogrammed CM had incomplete transcriptome remodeling. Wada (Wada et al., 2013) presented GATA4, MEF2C, TBX5, MESP1, and MYOCD as an alternative cocktail combination that generate more functional iCM. Moreover, Srivastava (Fu et al., 2013a) showed that the addition of MYOCD, ZFPM2, ESRRG, and MESP1 also work in generating human CM at higher efficiencies. Furthermore, Ding, et al. reported that this conversion could be achieved without the use of a viral delivery system (Cao et al., 2016a). Using small nine small molecules (CHIR99021, A83-01, BIX1294, AS8351, SC1, Y27632, OAC2, SU16F, and JNJ10198409) can reprogram human foreskin fibroblast into iCLMs with morphological, transcriptome, and epigenetic signatures similar to naïve human CM. Unfortunately, human reprogramming efficiencies are very low, in where the most efficient system can only achieve ~30% HFF reprogramming (Nam et al., 2013) as measured by Ca^{2+}

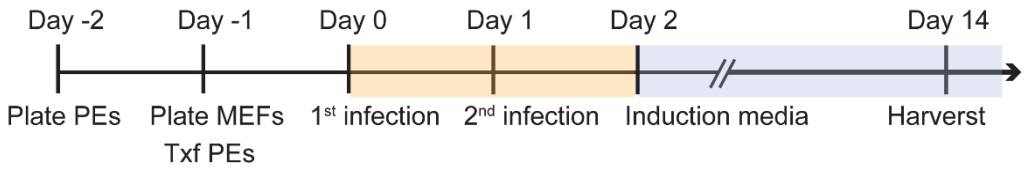
transients and beating frequency. Nonetheless, these studies represent a more reliable step towards clinical applications. Based on this body of work, it would be advantageous to set up an experimental pipeline that could capture hiPMs and perform detailed chromatin, methylomics, and single cell transcriptomics to uncover cell fate roadblocks and determinants of PM formation. Although this project faces major roadblocks, like the generation of novel PM reporter lines in primary human cells lines, successful results can allow for a platform to identify new drugs that will improve pacemaker function, study unique diseases, and discretely test individual TF contribution to disease.

COMPENDIUM OF OPTIMIZED PROTOCOLS

APENDIX A

Comprehensive protocol for the reprogramming of mouse embryonic fibroblasts towards cardiac

Procedure: Infection of mouse fibroblast (1x well of 6wlp→2x wells of 12wlp or 4x wells of 24wlp)



Day -2

1. Split Plat-E cells to 1e6 cells per well of a 6-well plate. Refer to **Table 1** for other formats.

Day -1

2a. Add 1ml of SureCoat to each well of a 6wlp. For IF, use 500µL per well of 24wlp with fibronectin coverslips.

2b. Fugene and Opti-MEM should be at room temperature before use. Transfect Plat-E cell using Fugene6 (Roche) system. Per sample **(6wlp format),** mix 6ul of Fugene in 60µl of Opti-MEM (Antibiotic free) in a 1.5ml eppendorf tube, gently mix by flicking, and incubate for 5 min at RT. Add Fugene:Opti-MEM mixture to DNA cocktail and mix thoroughly by flicking (Total DNA of 2µg). Incubate Fugene:DNA:Opti-MEM mixture for 15min at RT. Add transfection mix dropwise to the cells; mix well by gently rocking plate back and forward 4 times.

3. Incubate the transfected cells overnight at 37°C 5% CO_2. **Record time of transfection.**

4. Thaw fibroblasts and resuspend in 4ml of Fibroblast media. Wash @ 550 x *g* for 5min at 4ºC to remove excess DMSO. Plate at the given density from **Table 2** in fibroblast media and supplement with 10µM Y-27632 (Rock inhibitor).

Day 0 (Around the same time from step 3 ± 1hr)

5. To a 15ml conical tube, add polybrene to final concentration of 8µg/ml and mix well. 1:1000 from an 8mg/ml of a working stock.

6. 1st virus collection: 24hrs post transfection, filter the retroviral medium through a 0.45µm filter using a 6ml sterile disposable syringe into the 15ml conical tube with polybrene. Replace Plate-E dish with 2ml of fresh FB medium.

7. Aspirate the medium of the cultured fibroblast and add the freshly collected retroviral medium (It should be ~1.75ml of media from a well of a 6wlp); for a well of 12wlp use 750µL per well (x2) and for a well of a 24wlp use ~400µL per well (x4). Return to incubator and incubate overnight.

Day 1

Repeat step 5-7. Discard Plat-E cells after 2nd virus collection. **Checkpoint**: observe Plat-E GFP expression under the microscope (> %90); signal may be weak at this point.

Day 2

Change the medium from FB media to 18 media and replace every other day. Check MEFs GFP expression 48hrs post-2nd infection, and record observations. If using selection/small molecules, add antibiotic at this step

APENDIX B

Triple Immunofluorescence (mouse α-Actinin;rNPPA/rMyl2;chickenGFP)

1. Harvesting day: Carefully aspirate media of the plate.

2. Wash with 1x PBS, aspirate (1ml per well of a 6wlp).

3. Fix cells with 4% PFA (Dilute from 16% to 4% in H_2O) for 15 minutes at RT. For a well of a 24-well plate add 300µL.

Note! Fixed cells can be stored at 4°C up to a month in PBS before use.

4. Wash 3x with PBS-TX100 (0.1%) 5 minutes (300µL per well of a 24-well plate). Cover from light.

5. Block 10 minutes at RT with 300µL Universal Block Buffer 1X (Dilute from 10X).

6. Prepare 1ry ab mix. Refer to antibody table.

x1 sample	x4 samples
105µL 1x PBS	420µL 1x PBS
105µL 1x Blocking	420µL 1x Blocking
0.525µL per 1ab ((m-α-actinin/rNPPA-rMyl2/cGFP)	2.1uµL per 1ab (m-αActinin/rNPPA-rMyl2/cGFP)
200µL per well (24wlp)	200µL per well (24wlp)

7. The next day wash x3 with 300µL of PBST for 5 minutes at RT.

8. During the last wash prepare 2ry ab mix (1:400). *Alexa* fluorochromes*:*

x1 sample	x4 samples
105µL 1x PBS	420µL 1x PBS
105µL 1x Blocking	4420µL 1x Blocking
0.3µL of 2ry ab (cA488/mA555/rA647)	1.2µL of 2ab (cA488/mA555/rA647)
200µL per well (24wlp)	200µL per well (24wlp)

Key point: Make sure to protect samples from light

9. Wash x3 with 1ml of 0.1% PBST for 5 minutes.

10. Mount cover slip onto glass slide with Vectashield/DAPI (2.5ul is sufficient)

11. Seal with Wet & Wild nail polish.

    Sealed slides can be stored at 4°C up to a month before analysis and quantification under the microscope.

12. Analyze by confocal/fluorescence microscopy

APENDIX C

Intracellular FACS staining (cTnT and GFP)

1. Detach cells with 0.25% Trypsin for ~2-5mins. (500µL per well of a 6wlp)
2. Collect cells with PBS w/ 5% FBS (4ml per well of 6wlp). Perform twice to recover maximum # cells.
3. Pipet up and down to resuspend. Collect cells by centrifuging for 5min at 500 x *g* 4°C. Aspirate supernatant.
4. Fix and permeabilize with BD Fix/Perm buffer in 0.2ml/1e6 cells in 1.5 ml eppendorf tube for 15 mins on ice.
5. Wash 1x: add 1ml BD perm/wash buffer.
6. Collect cells: Centrifuge 5min at 500 x *g*; store at 4°C or continue with the staining protocol. Carefully aspirate supernatant and tap tube to loosen pellet with remaining solution.
7. Add 1ry in 50µL BD Perm/Wash buffer, incubate for 1-2hrs at RT **(m-TnT 1:100, rabbit-GFP 1:100)**

x1 sample	**x8 samples**
51µL 1x Perm/Wash Buffer	408µL 1x PBS
0.51µL of 1ry ab (m-cTnT/rGFP)	4.08µL of 1ry ab (m-cTnT/rGFP)
50µL per sample	50µL per sample

8. Add 1ml cold BD perm/wash buffer and tap tube to gently mix.
9. Collect cells; 5 min @ 500 x *g*.
10. Incubate for 60 mins with 2ry in 50µL BD perm/wash buffer at RT (α-Rabbit A488 1:200, Anti-Mouse A647 1:200 for cTnT). Protect from light.

x1 sample	x8 samples
551µL 1x Perm/Wash Buffer	408µL 1x PBS
0.26µL of 2ry ab (mA647/rA488)	2.08µL of 2ry ab (mA647/rA488)
50µL per sample	50µL per sample

11. Wash with 1ml of ice cold BD perm/wash buffer.

12. Collect cells: 5 mins @ 500 x *g* RT.

13. Add 350µL of staining buffer (depends on pellet size; usual range is from 250-400µL)

14. Store at 4°C until ready for analysis.

Notes:

For A647 use FL-4 in the FACS calibur, and FL-1 for A488.

X-axis: Hcn4-GFP (If HCN4 MEFs were used). Y-axis: cTnT.

Use 350µL of FACs buffer to avoid clogging of the cytometer.

APENDIX D

Optimized protocol for the co-Immunoprecipitation of Hand2 and Tbx5

Day 1

1. Wash cells twice with 1X PBS and add in 1 mL (6-cm plate) of **AFP_IP lysis buffer** with inhibitors.
2. Lyse cells in dish, and transfer into microcentrifuge tubes when all the cells are visibly removed.

   Note! Sonication is not recommended for Co-IP.

3. Triturate cell mixture on ice with 0.3 mL insulin syringe (30g). Do it 10-20 times. Until your hand hurts.
4. Incubate on ice for 30 min.
3. Centrifuge at 10,000 x *g* for 10 min at 4°C to pellet cell debris. Transfer ~900 µL S/N to new tube.
4. Remove 500 µL of S/N and add 500 µL of **AFP_IP lysis buffer** IP lysate!
4. Leftover ~400 µL of S/N is your Input sample! Free at the -80C until ready for IB.
5. Add 2 µg (or 2 µL) of antibody for IP to S/N protein and incubate overnight at 4°C with rotation.

   Note! Include a control IP with IgG or an unrelated antibody.

Day 2

7. Vortex Dynabeads and take 50 µL/reaction. Wash x3 with 200 µL of AFP_IP lysis buffer with magnetic sorter.
8. Resuspend in 50 µL of AFP_IP lysis buffer per IP (use 10% excess; 100 µL beads in 110 µL buffer).
8. Add 50 µL of beads to each IP reaction and mix by pipetting on ice.
9. Incubate 1-**2** hrs at 4°C rotating.
10. Apply magnet and discard supernatant.

    Control! Keep S/N for troubleshooting.

11. Wash x5 with 500 µL IP lysis buffer, with 5 min rotating. Perform in this in the cold room.

    Control! Keep 1st wash for troubleshooting.

12. On the last wash, transfer the whole volume to a clean tube. To reduce non-specific binding.

13. Apply magnet and discard all supernatant.

14. Resuspend beads with 25 µL of 100 mM glycine, pH 2.0.

    Note! For a stronger elution, apply 1X Laemmli buffer directly, heat at 95°C for 10 min and load.

15. Incubate 10 min at RT or 55°C (stronger). Place on magnet, then transfer the S/N to a clean tube.

    Control! Keep beads for troubleshooting.

16. Neutralize the protein with 10 µL of 1M Tris, pH 8.0.

17. Add 35 µL of 2x Laemmli loading buffer with β-ME.

19. Heat at 95°C for 10 min and load.

APENDIX E

Optimized Omni-ATAC protocol for low cell samples

Buffers and Reagents:

Tn5 obtained from the Hon lab. Keep stocks at -80°C. Working stocks store at -20°C

Zymo DNA Clean and concentrator (Cat # D4014).

KAPA library quantification kit (Cat # KK4828).

ATAC (RSB)

Reagent	[Final]	for 50 mL	10 mL	2 mL
1M Tris-HCL pH 7.4	10 mM	500 µl	100 µl	20 µl
5M NaCl	10 mM	100 µl	20 µl	4 µl
1M $MgCl_2$	3 mM	150 µl	30 µl	6 µl
Sterile H_20	NA	49.25mL	9.85 mL	1.97 mL

Detergents (All detergents will be considered as 100X stock solutions)

Digitonin (Sigma D141-100MG). Make a **1% (w/v)** in DMSO (10mg/mL). Can be stored at -20C for up to 6 months if minimal (<5) thaw cycles are used.

Tween-20 (GSO 5336 100%). Dilute in H_2O to make a **10% stock** and store at 4C. This will be 100X stock.

NP40. Dilute in H_2O to make a **10% stock** and store at 4°C. This will be 100X stock.

2x TD buffer

Reagent	[Final]	for 100 mL	1mL	100µL
1M Tris-HCL pH 7.6	20 mM	2 mL	20 µl	2 µl
1M $MgCl_2$	10 mM	1 mL	10 µl	1 µl
*DMF @ 4°C Deli	20%	20 mL	200 µl	20 µl
Sterile H_20	NA	Bring to 100 mL	770 µl	77 µl

A. Transposition reaction

1. Pellet 12.5k viable MEFs at 500 g at 4°C for 5 mins in a fixed angle centrifuge.
2. Remove supernatant, avoiding the *visible* pellet, using two pipetting steps. Remove to 100 µl with a P1000, then the rest with a P200.
3. Add 50 µl of cold **RSB** with **0.1% NP40, 0.1% Tween-20, and 0.01% Digitonin** and pipette up and down 3 times. (1µl detergent per 100µl of RSB).
4. Incubate on ice for **3 mins**.
5. Wash out lysis with 1 mL of cold **RSB** containing **0.1% Tween-20** but No NP40 or Digitonin and invert tube 3 times to mix.
6. Pellet nuclei at 500 *g* for 10 mins at 4°C in a fixed angle centrifuge.
7. Remove all of the supernatant, avoiding the *visible* pellet, using two pipetting steps. Remove to 100 µl with a P1000, then the rest with a P200.

8. **Note!** Start here if you are coming from SB nuclei isolation protocol.
   Resuspend cell pellet in **50 µl** of transposition mixture by pipetting 6 times.

Transposition mix				
Reagent	**1x**	**2.5x**	**4.5x**	**20.5x**
2X-TD buffer	25 µl	62.50 µl	112.5 µl	512.5 µl
Tn5 (100nM Final)	2.5 µl	6.25 µl	11.25 µl	51.25 µl
1X PBS	16.5 µl	41.25 µl	74.25 µl	338.25 µl
1% Digitonin	0.5 µl	1.25 µl	2.25 µl	10.25 µl
10% Tween-20	0.5 µl	1.25 µl	2.25 µl	10.25 µl
H_20	5 µl	12.50 µl	22.5 µl	102.5 µl

9. Incubate reaction at 37°C for 30 mins in a thermomixer with 1000 RPM mixing.
10. Cleanup with Zymo DNA Clean and concentrator-5 kit.
    a. Add 250 µl of **DNA Binding Buffer** and mix briefly by vortexing.
    b. Transfer mixture to Zymo-Spin column in a collection tube.
    c. Centrifuge for 30 sec at 17900 *g*. Discard flow through.
    d. Add 200 µL **DNA Wash buffer** to the column. Centrifuge for 30 secs. Discard flow through. Repeat wash, but do it for 2 min to eliminate EtOH.
    e. Transfer column to a labeled Eppendorf tube. Add ~21 µL of **DNA elution buffer**. Incubate for 1min at RT. Centrifuge for 1 min at 17900 *g*.
    f. Eluted DNA in (~20 µl) can store at -20°C.
    g.

B. Pre-amplification and determination of additional cycles (22 mins).

With Omni-ATAC, many libraries require little to no additional cycles to be performed. However, a minimal of 5 pre-amplification cycles are required to add the Illumina adaptor sequences.

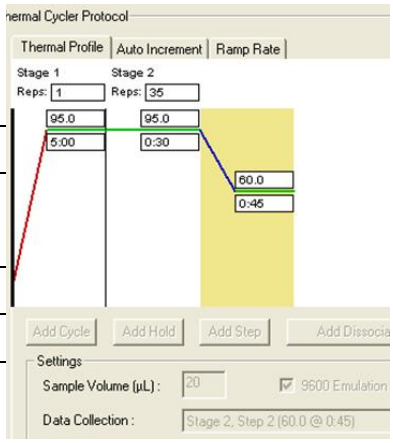
PCR for Illumina sequencing		
Reagent	**1x**	**4.5x**
HiFi Buffer (5X)	10 µL	45 µL
dNTP (10 mM)	1.5 µL	6.75 µL
H_2O	12.5 µL	56.25 µL
KAPA HiFi Enzyme	1 µL	4.5 µL
DNA (From Zymo)	20 µL	NA
Ad1.noMX [10 µM]	2.5 µL	11.25 µL
Ad2.X [10 µM]	2.5 µL	2.5 ea

Pre-amplification

1. Initial PCR (Bio-Rad **ATAC-I**)
   a. Pre-PCR hold: 72°C for 5 min
   b. Initial denaturation: 98°C for 30secs
   c. Cycle (5 cycles)

i. Denaturation: 98°C for 10secs
ii. Annealing: 63°C for 30secs
iii. Extension: 72°C for 1min.

2. Dilute 1 µl of the "pre-amplification" reaction onto 999 µl of Molecular grade H_2O and use KAPA library quantification kit to estimate number of cycles to amplify. Refer to excel sheet (Kapa ATAC) for template and see below for qPCR set-up



Reagent	x1	x26		
2x Master Mix + Primer mix	12 µl	312 µl		
H_20	4 µl	104 µl		
DNA	4 µl	NA		

a. Total volume 20 µl per well.
b. Run the DNA standards and the experimental samples in duplicate. If possible do in triplicate.
c. Only use 4 µl of the 1:1000 dilution of the experimental samples. Don't dilute the DNA standards and use 4 µl directly.

3. Use the remainder of the pre-amplified DNA (49 µl) , and run the required cycles based on the template on (Bio-Rad **ATAC-II**):
   a. Initial denaturation: 98°C for 30secs
   b. Cycle (X cycles from qPCR)
      i. Denaturation: 98°C for 10secs
      ii. Annealing: 63°C for 30secs
      iii. Extension: 72°C for 1min.

C. Post-amplification cleanup

1. Warm AMPure XP beads to RT for at least 30 min before use. Calculate original volume from step B.3.
2. Add **0.5x volume** of RT AMPure XP beads (i.e. 22 µl for 44 µl original volume).
   a. Gently pipette up and down 10x to mix with P200.
3. Incubate at RT for 15 minutes.
4. Place epi tubes in magnetic rack for 5 minutes at RT making sure the liquid clears.

5. Transfer supernatant to new labeled tube. Make sure not to disturb the beads. It is ok to leave behind 1-2 µl of volume. Discard the beads.
6. Add 1.2x RT AMPure XP beads (Based on new volume) to Supernatant.
    a. Pipette up and down 10x to mix with P200.
7. Incubate at RT for 10 minutes.
8. Place epi tubes in magnetic rack for 5 minutes at RT.
9. Discard supernatant.
10. Wash beads with 200 µl of freshly prepared 80% EtOH. Incubate for 30 secs at RT. Discard EtOH wash with P200. **Note!** Do not disturb beads.
11. Wash beads with 200 µl of freshly prepared 80% EtOH. Incubate for 30 secs at RT. Discard EtOH wash with P200. **Note!** Do not disturb beads.
12. Air dry for 15 mins. Note: Make sure to remove all of the EtOH. Carryover will significantly affect sequencing.
13. Add 20 µl of nuclease free H_2O. Remove tubes from magnet.
    a. Pipette up and down 10x to mix with a P200.
    b. Incubate at 55°C for 5 mins at 1000 RPMs.
    c. Place epi tubes in magnetic rack for 5 minutes at RT.
    d. Transfer the solution to a new pre-labeled tube.

D. Qubit quantification

Set up master mix as follows:

Reagent	Volume	For x6 rxns
Qubit dsDNA HS Buffer	200 µl	1.2 mL
Dye High sensitivity	1 µl	6 µl

For the standards:

190 µl M.M + 10 µl of Standard sample into a 500 µl Eppendorf tube.

For the experimental samples:

199 µl M.M + 1 µl of experimental sample into a 500 µl Eppendorf tube.

1. Vortex
2. Incubate 2 mins at RT
3. Prep Qubit machine (Touch screen; close lid before reading each sample)
    a. First S1
    b. Second S2

c. Then experimental samples.

4. Record each concentration.

E. TapeStation quantification (D1000)

1. Equilibrate reagents at room temperature.
2. Use PCR tubes without caps.
3. Take 3 µl of D1000 Sample Buffer.
4. 1 µl of your DNA sample.
    a. If using D1000 ladder. Dilute 1 µl of ladder into 3 µl of sample buffer.
5. Vortex PCR tubes at 2000 rpm for 1 min.
6. Quick-spin to pellet down.
7. Add tape, with barcode facing to the right.
8. Empty tip box in station and add full row of new tips.
9. Add PCR tubes into correct location (without caps). Designate location in the computer program.
10. Add ladder if using uneven number of samples, or use artificial ladder for the D1000 tape.
11. Click Start. (It takes around 1.5 min per sample). **Note!** Save unused tips and place back into tip box.

D. Sequencing

1. Species: mouse
2. Library prep: ATAC-seq.
3. Project size: Usually = “My samples will fill a flowcell”
4. Sequencing run request: NextSeq 500 High Output (400 M)
5. Run 1: (42)
6. Run 2: (42)
7. Index Read 1 (i7): 8
8. Index Read 2 (i5): 0

APENDIX F

Supplementary Table 1: Oligos for PCR (*See Nature. Buenrostro, 2013*)

Greenleaf primers ATAC-seq

Index ID		**Adapter 8bp**
Ad1_noMX:	AATGATACGGCGACCACCGAGATCTACACTCGTCGGCAGCGTCAGATGTG	None
Ad2.1	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGGAGATGT	TAAGGCGA
Ad2.2	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGGGCTCGGAGATGT	CGTACTAG
Ad2.3	CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGGCTCGGAGATGT	AGGCAGAA
Ad2.4	CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTCTCGTGGGCTCGGAGATGT	TCCTGAGC
Ad2.5	CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTCTCGTGGGCTCGGAGATGT	GGACTCCT
Ad2.6	CAAGCAGAAGACGGCATACGAGATCATGCCTAGTCTCGTGGGCTCGGAGATGT	TAGGCATG
Ad2.7	CAAGCAGAAGACGGCATACGAGATGTAGAGAGGTCTCGTGGGCTCGGAGATGT	CTCTCTAC
Ad2.8	CAAGCAGAAGACGGCATACGAGATCCTCTCTGGTCTCGTGGGCTCGGAGATGT	CAGAGAGG
Ad2.9	CAAGCAGAAGACGGCATACGAGATAGCGTAGCGTCTCGTGGGCTCGGAGATGT	GCTACGCT
Ad2.10	CAAGCAGAAGACGGCATACGAGATCAGCCTCGGTCTCGTGGGCTCGGAGATGT	CGAGGCTG
Ad2.11	CAAGCAGAAGACGGCATACGAGATTGCCTCTTGTCTCGTGGGCTCGGAGATGT	AAGAGGCA
Ad2.12	CAAGCAGAAGACGGCATACGAGATTCCTCTACGTCTCGTGGGCTCGGAGATGT	GTAGAGGA
Ad2.13	CAAGCAGAAGACGGCATACGAGATATCACGACGTCTCGTGGGCTCGGAGATGT	GTCGTGAT
Ad2.14	CAAGCAGAAGACGGCATACGAGATACAGTGGTGTCTCGTGGGCTCGGAGATGT	ACCACTGT
Ad2.15	CAAGCAGAAGACGGCATACGAGATCAGATCCAGTCTCGTGGGCTCGGAGATGT	TGGATCTG
Ad2.16	CAAGCAGAAGACGGCATACGAGATACAAACGGGTCTCGTGGGCTCGGAGATGT	CCGTTTGT
Ad2.17	CAAGCAGAAGACGGCATACGAGATACCCAGCAGTCTCGTGGGCTCGGAGATGT	TGCTGGGT
Ad2.18	CAAGCAGAAGACGGCATACGAGATAACCCCTCGTCTCGTGGGCTCGGAGATGT	GAGGGGTT
Ad2.19	CAAGCAGAAGACGGCATACGAGATCCCAACCTGTCTCGTGGGCTCGGAGATGT	AGGTTGGG
Ad2.20	CAAGCAGAAGACGGCATACGAGATCACCACACGTCTCGTGGGCTCGGAGATGT	GTGTGGTG
Ad2.21	CAAGCAGAAGACGGCATACGAGATGAAACCCAGTCTCGTGGGCTCGGAGATGT	TGGGTTTC
Ad2.22	CAAGCAGAAGACGGCATACGAGATTGTGACCAGTCTCGTGGGCTCGGAGATGT	TGGTCACA
Ad2.23	CAAGCAGAAGACGGCATACGAGATAGGGTCAAGTCTCGTGGGCTCGGAGATGT	TTGACCCT
Ad2.24	CAAGCAGAAGACGGCATACGAGATAGGAGTGGGTCTCGTGGGCTCGGAGATGT	CCACTCCT

APENDIX G

Scripts for Rstudio

GoBarplots: For the generation of GO terms in barplot format

```
#Load libraries

library(ggplot2)

library(scales)

library(RColorBrewer)

library(forecast)

#Set up matrix data

go_data <- read.csv("Title of csv.csv", header = TRUE)

go_data_mat <- go_data[1:5,
c("Description","q.value","Enrichment","Adj.p.value")]

go_data_mat$Description <- factor(go_data_mat$Description, levels =
go_data_mat$Description[order(go_data_mat$Adj.p.value)]) #This orders the
data in ascending order


p <-ggplot(go_data_mat,

          aes(x= factor(Description), y= Adj.p.value,
fill=as.factor(Enrichment))) +

  geom_bar(stat = "identity", width = 0.6, position =
position_dodge(width=1)) +

  scale_fill_brewer(palette = "Reds")

p + coord_flip()+

  theme_light()+

  scale_y_continuous(expand = c(0,0))

ggsave(file="Title of desired file.pdf", dpi = 600, width = 8, height = 2,
units = "in") #12 for med size letters
```

pHeatmap Script: For the generation of mini Heatmaps

```
#Load libraries

library(pheatmap)

library(RColorBrewer)
```

```
#Set up matrix data

mat_data <- read.csv("Title of csv.csv", header = T)

mat_data_test <- mat_data[1:9,] #Brackets tell the script to only use the
first 500 entries, Make sure to edit to your dataset

rownames(mat_data_test) <- mat_data_test[,1] #The ,1 tells the rownames are
the 1st column

hmcols<- colorRampPalette(c("blue","gold"))(256)

pheatmap(mat_data_test[1:9,2:7], #First delimitation sets how many to use on
the list ## Second, what columns to use 2nd and 3rd

         scale="row",

         #brakes=bk,

         kmeans_k=NA,

         cluster_rows=FALSE,

         cluster_cols=FALSE,

         col=hmcols,

         #color=brewer.pal(9, "YlGnBu"),

         treeheight_row=0,

         treeheight_col=0,

         legend=TRUE,

         show_colnames=TRUE,

         show_rownames =TRUE,

         fontsize=10,

         fontface="plain",

         border_color="white",

         width=11, heigth=8, cellheight = 10, cellwidth=10)

pdf()
```

BIBLIOGRAPHY

Abad, M., Hashimoto, H., Zhou, H., Morales, M.G., Chen, B., Bassel-Duby, R., and Olson, E.N. (2017a). Notch Inhibition Enhances Cardiac Reprogramming by Increasing MEF2C Transcriptional Activity. Stem cell reports *8*, 548-560.

Abad, M., Hashimoto, H., Zhou, H., Morales, M.G., Chen, B., Bassel-Duby, R., and Olson, E.N.J.S.C.R. (2017b). Notch inhibition enhances cardiac reprogramming by increasing MEF2C transcriptional activity. *8*, 548-560.

Addis, R.C., and Epstein, J.A.J.N.m. (2013). Induced regeneration—the progress and promise of direct reprogramming for heart repair. *19*, 829.

Addis, R.C., Ifkovits, J.L., Pinto, F., Kellam, L.D., Esteso, P., Rentschler, S., Christoforou, N., Epstein, J.A., Gearhart, J.D.J.J.o.m., and cardiology, c. (2013). Optimization of direct fibroblast reprogramming to cardiomyocytes using calcium activity as a functional measure of success. *60*, 97-106.

Ahmed, Z., and Ucar, D. (2017). I-ATAC: interactive pipeline for the management and pre-processing of ATAC-seq samples. PeerJ *5*, e4040.

Albrecht, A., and Mundlos, S. (2005). The other trinucleotide repeat: polyalanine expansion disorders. Curr Opin Genet Dev *15*, 285-293.

Ali, S.R., Hippenmeyer, S., Saadat, L.V., Luo, L., Weissman, I.L., and Ardehali, R. (2014). Existing cardiomyocytes generate cardiomyocytes at a low rate after birth in mice. Proc Natl Acad Sci U S A *111*, 8850-8855.

Anderson, R.H., Brown, N.A., and Moorman, A.F.J.D.d.a.o.p.o.t.A.A.o.A. (2006). Development and structures of the venous pole of the heart. *235*, 2-9.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.

Anwar, K., Han, T., and Kim, S.M. (2011). Reversible sealing techniques for microdevice applications. Sensors and Actuators B: Chemical *153*, 301-311.

Aoki, H., Sadoshima, J., and Izumo, S.J.N.m. (2000). Myosin light chain kinase mediates sarcomere organization during cardiac hypertrophy in vitro. *6*, 183.

Aronesty, E. (2013). Comparison of sequencing utility programs. The Open Bioinformatics Journal *7*.

Atkinson, A., Inada, S., Li, J., Tellez, J.O., Yanni, J., Sleiman, R., Allah, E.A., Anderson, R.H., Zhang, H., Boyett, M.R.*, et al.* (2011). Anatomical and molecular mapping of the left and right ventricular His-Purkinje conduction networks. J Mol Cell Cardiol *51*, 689-701.

Baddeley, D., Crossman, D., Rossberger, S., Cheyne, J.E., Montgomery, J.M., Jayasinghe, I.D., Cremer, C., Cannell, M.B., and Soeller, C. (2011). 4D super-resolution microscopy with conventional fluorophores and single wavelength excitation in optically thick cells and tissues. PloS one *6*, e20645.

Barnes, R.M., Firulli, B.A., VanDusen, N.J., Morikawa, Y., Conway, S.J., Cserjesi, P., Vincentz, J.W., and Firulli, A.B.J.C.r. (2011). Hand2 loss-of-function in Hand1-expressing cells reveals distinct roles in epicardial and coronary vessel development. *108*, 940-949.

Barron, F., Woods, C., Kuhn, K., Bishop, J., Howard, M.J., and Clouthier, D.E.J.D. (2011). Downregulation of Dlx5 and Dlx6 expression by Hand2 is essential for initiation of tongue morphogenesis. *138*, 2249-2259.

Bass, G.T., Ryall, K.A., Katikapalli, A., Taylor, B.E., Dang, S.T., Acton, S.T., and Saucerman, J.J. (2012a). Automated image analysis identifies signaling pathways regulating distinct signatures of cardiac myocyte hypertrophy. J Mol Cell Cardiol *52*, 923-930.

Bass, G.T., Ryall, K.A., Katikapalli, A., Taylor, B.E., Dang, S.T., Acton, S.T., Saucerman, J.J.J.J.o.m., and cardiology, c. (2012b). Automated image analysis identifies signaling pathways regulating distinct signatures of cardiac myocyte hypertrophy. *52*, 923-930.

Bergmann, O., Bhardwaj, R.D., Bernard, S., Zdunek, S., Barnabe-Heider, F., Walsh, S., Zupicich, J., Alkass, K., Buchholz, B.A., Druid, H.*, et al.* (2009). Evidence for cardiomyocyte renewal in humans. Science *324*, 98-102.

Bergmann, O., Zdunek, S., Alkass, K., Druid, H., Bernard, S., and Frisén, J.J.E.c.r. (2011). Identification of cardiomyocyte nuclei and assessment of ploidy for the analysis of cell turnover. *317*, 188-194.

Bhatia, S.N., and Ingber, D.E. (2014). Microfluidic organs-on-chips. Nat Biotechnol *32*, 760-772.

Blazek, A.D., Paleo, B.J., and Weisleder, N. (2015). Plasma Membrane Repair: A Central Process for Maintaining Cellular Homeostasis. Physiology *30*, 438-448.

Bootman, M.D., Smyrnias, I., Thul, R., Coombes, S., and Roderick, H.L. (2011). Atrial cardiomyocyte calcium signalling. Biochim Biophys Acta *1813*, 922-934.

Bray, M.A., Sheehy, S.P., Parker, K.K.J.C.m., and cytoskeleton, t. (2008). Sarcomere alignment is regulated by myocyte shape. *65*, 641-651.

Bruneau, B.G.J.N. (2008). The developmental genetics of congenital heart disease. *451*, 943.

Buckingham, M., Meilhac, S., and Zaffran, S.J.N.R.G. (2005). Building the mammalian heart from two sources of myocardial cells. *6*, 826.

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. Nat Methods *10*, 1213-1218.

Buganim, Y., Faddah, D.A., Cheng, A.W., Itskovich, E., Markoulaki, S., Ganz, K., Klemm, S.L., van Oudenaarden, A., and Jaenisch, R. (2012). Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell *150*, 1209-1222.

Burns, J.C., Friedmann, T., Driever, W., Burrascano, M., and Yee, J.K. (1993). Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. Proc Natl Acad Sci U S A *90*, 8033-8037.

Cai, D., and Neyer, A. (2010). Cost-effective and reliable sealing method for PDMS (PolyDiMethylSiloxane)-based microfluidic devices with various substrates. Microfluidics and Nanofluidics *9*, 855-864.

Cai, Z., Qiu, W., Shao, G., and Wang, W. (2013). A new fabrication method for all-PDMS waveguides. Sensors and Actuators A: Physical *204*, 44-47.

Cao, N., Huang, Y., Zheng, J., Spencer, C.I., Zhang, Y., Fu, J.-D., Nie, B., Xie, M., Zhang, M., Wang, H.*, et al.* (2016a). Conversion of human fibroblasts into functional cardiomyocytes by small molecules. Science *352*, 1216-1220.

Cao, N., Huang, Y., Zheng, J., Spencer, C.I., Zhang, Y., Fu, J.-D., Nie, B., Xie, M., Zhang, M., and Wang, H.J.S. (2016b). Conversion of human fibroblasts into functional cardiomyocytes by small molecules. *352*, 1216-1220.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., and Moffat, J.J.G.b. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *7*, R100.

Cerrone, M., Montnach, J., Lin, X., Zhao, Y.T., Zhang, M., Agullo-Pascual, E., Leo-Macias, A., Alvarado, F.J., Dolgalev, I., Karathanos, T.V.*, et al.* (2017). Plakophilin-2 is required for transcription of genes that control calcium cycling and cardiac rhythm. Nat Commun *8*, 106.

Chanda, S., Ang, C.E., Davila, J., Pak, C., Mall, M., Lee, Q.Y., Ahlenius, H., Jung, S.W., Südhof, T.C., and Wernig, M.J.S.c.r. (2014). Generation of induced neuronal cells by the single reprogramming factor ASCL1. *3*, 282-296.

Chen, Y., Yang, Z., Zhao, Z.-A., Shen, Z.J.S.c.r., and therapy (2017). Direct reprogramming of fibroblasts into cardiomyocytes. *8*, 118.

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. Genes Dev *8*, 2293-2301.

Chong, J.J., Yang, X., Don, C.W., Minami, E., Liu, Y.-W., Weyers, J.J., Mahoney, W.M., Van Biber, B., Cook, S.M., and Palpant, N.J.J.N. (2014). Human embryonic-stem-cell-derived cardiomyocytes regenerate non-human primate hearts. *510*, 273.

Chopra, A., Patel, A., Shieh, A.C., Janmey, P.A., and Kresh, J.Y.J.P.o. (2012). α-Catenin localization and sarcomere self-organization on N-cadherin adhesive patterns are myocyte contractility driven. *7*, e47592.

Christoffels, V.M., Mommersteeg, M.T., Trowe, M.-O., Prall, O.W., de Gier-de Vries, C., Soufan, A.T., Bussen, M., Schuster-Gossler, K., Harvey, R.P., and Moorman, A.F.J.C.r. (2006). Formation of the venous pole of the heart from an Nkx2–5–negative precursor population requires Tbx18. *98*, 1555-1563.

Christoforou, N., Chellappan, M., Adler, A.F., Kirkton, R.D., Wu, T., Addis, R.C., Bursac, N., and Leong, K.W.J.P.o. (2013). Transcription factors MYOCD, SRF, Mesp1 and SMARCD3 enhance the cardio-inducing effect of GATA4, TBX5, and MEF2C during direct cellular reprogramming. *8*, e63577.

Coffin, J.M., Hughes, S.H., and Varmus, H.E. (1997). The Interactions of Retroviruses and their Hosts. In Retroviruses, J.M. Coffin, S.H. Hughes, and H.E. Varmus, eds. (Cold Spring Harbor (NY)).

Connelly, J.T., Gautrot, J.E., Trappmann, B., Tan, D.W., Donati, G., Huck, W.T., and Watt, F.M. (2010). Actin and serum response factor transduce physical cues from the microenvironment to regulate epidermal stem cell fate decisions. Nat Cell Biol *12*, 711-718.

Conner, D.A. (2001). Mouse Embryo Fibroblast (MEF) Feeder Cell Preparation. In Current Protocols in Molecular Biology (John Wiley & Sons, Inc.).

Conway, S.J., Firulli, B., and Firulli, A.B. (2010). A bHLH code for cardiac morphogenesis. Pediatr Cardiol *31*, 318-324.

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. Nat Methods *14*, 959-962.

Corrado, D., Basso, C., and Judge, D.P. (2017). Arrhythmogenic Cardiomyopathy. Circ Res *121*, 784-802.

Dai, Y.S., and Cserjesi, P. (2002). The basic helix-loop-helix factor, HAND2, functions as a transcriptional activator by binding to E-boxes as a heterodimer. The Journal of biological chemistry *277*, 12604-12612.

Davis, R.L., Weintraub, H., and Lassar, A.B.J.C. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *51*, 987-1000.

Dirkx, E., Gladka, M.M., Philippen, L.E., Armand, A.-S., Kinet, V., Leptidis, S., El Azzouzi, H., Salic, K., Bourajjaj, M., and Da Silva, G.J.J.N.c.b. (2013). Nfat and miR-25 cooperate to reactivate the transcription factor Hand2 in heart failure. *15*, 1282.

Dunbar, C.E., High, K.A., Joung, J.K., Kohn, D.B., Ozawa, K., and Sadelain, M. (2018). Gene therapy comes of age. Science *359*.

Efe, J.A., Hilcove, S., Kim, J., Zhou, H., Ouyang, K., Wang, G., Chen, J., and Ding, S.J.N.c.b. (2011). Conversion of mouse fibroblasts into cardiomyocytes using a direct reprogramming strategy. *13*, 215.

Ephrussi, A., Church, G.M., Tonegawa, S., and Gilbert, W.J.S. (1985). B lineage--specific interactions of an immunoglobulin enhancer with cellular factors in vivo. *227*, 134-140.

Espinoza-Lewis, R.A., Yu, L., He, F., Liu, H., Tang, R., Shi, J., Sun, X., Martin, J.F., Wang, D., Yang, J.*, et al.* (2009). Shox2 is essential for the differentiation of cardiac pacemaker cells by repressing Nkx2-5. Dev Biol *327*, 376-385.

Fawcett, J.M., Harrison, S.M., and Orchard, C.H. (1998). A method for reversible permeabilization of isolated rat ventricular myocytes. Exp Physiol *83*, 293-303.

Fellmann, C., Gowen, B.G., Lin, P.C., Doudna, J.A., and Corn, J.E. (2017). Cornerstones of CRISPR-Cas in drug discovery and therapy. Nat Rev Drug Discov *16*, 89-100.

Fernandez-Perez, A., and Munshi, N.V. (2017). Assessing Cardiomyocyte Subtypes Following Transcription Factor-mediated Reprogramming of Mouse Embryonic Fibroblasts. J Vis Exp.

Firulli, A.B., Firulli, B.A., Wang, J., Rogers, R.H., and Conway, S.J. (2010a). Gene replacement strategies to test the functional redundancy of basic helix-loop-helix transcription factor. Pediatr Cardiol *31*, 438-448.

Firulli, A.B., Firulli, B.A., Wang, J., Rogers, R.H., and Conway, S.J.J.P.c. (2010b). Gene Replacement Strategies to Test the Functional Redundancy of Basic Helix–Loop–Helix Transcription Factor. *31*, 438-448.

Firulli, A.B., McFadden, D.G., Lin, Q., Srivastava, D., and Olson, E.N. (1998). Heart and extra-embryonic mesodermal defects in mouse embryos lacking the bHLH transcription factor Hand1. Nat Genet *18*, 266-270.

Firulli, B.A., Hadzic, D.B., McDaid, J.R., and Firulli, A.B.J.J.o.B.C. (2000). The Basic Helix-Loop-Helix Transcription factorsdHAND and eHAND Exhibit Dimerization

Characteristics That Suggest Complex Regulation of Function. *275*, 33567-33573.

Fogel, I., and Sagi, D.J.B.c. (1989). Gabor filters as texture discriminator. *61*, 103-113.

Fu, J.-D., Stone, N.R., Liu, L., Spencer, C.I., Qian, L., Hayashi, Y., Delgado-Olguin, P., Ding, S., Bruneau, B.G., and Srivastava, D.J.S.c.r. (2013a). Direct reprogramming of human fibroblasts toward a cardiomyocyte-like state. *1*, 235-247.

Fu, J.D., Stone, N.R., Liu, L., Spencer, C.I., Qian, L., Hayashi, Y., Delgado-Olguin, P., Ding, S., Bruneau, B.G., and Srivastava, D. (2013b). Direct reprogramming of human fibroblasts toward a cardiomyocyte-like state. Stem Cell Reports *1*, 235-247.

Fu, Y., Huang, C., Xu, X., Gu, H., Ye, Y., Jiang, C., Qiu, Z., and Xie, X.J.C.r. (2015). Direct reprogramming of mouse fibroblasts into cardiomyocytes with chemical cocktails. *25*, 1013.

Galli, A., Robay, D., Osterwalder, M., Bao, X., Bénazet, J.-D., Tariq, M., Paro, R., Mackem, S., and Zeller, R.J.P.g. (2010). Distinct roles of Hand2 in initiating polarity and posterior Shh expression during the onset of mouse limb bud development. *6*, e1000901.

Gao, X., Tao, Y., Lamas, V., Huang, M., Yeh, W.H., Pan, B., Hu, Y.J., Hu, J.H., Thompson, D.B., Shu, Y.*, et al.* (2018). Treatment of autosomal dominant hearing loss by in vivo delivery of genome editing agents. Nature *553*, 217-221.

Garg, V., Kathiriya, I.S., Barnes, R., Schluterman, M.K., King, I.N., Butler, C.A., Rothrock, C.R., Eapen, R.S., Hirayama-Yamada, K., Joo, K.*, et al.* (2003). GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. Nature *424*, 443-447.

George, R.M., and Firulli, A.B. (2018). Hand Factors in Cardiac Development. Anat Rec (Hoboken).

Gervais, T., El-Ali, J., Gunther, A., and Jensen, K.F. (2006). Flow-induced deformation of shallow microfluidic channels. Lab on a chip *6*, 500-507.

Go, A.S., Mozaffarian, D., Roger, V.L., Benjamin, E.J., Berry, J.D., Blaha, M.J., Dai, S., Ford, E.S., Fox, C.S., and Franco, S.J.C. (2014). Executive summary: heart disease and stroke statistics—2014 update: a report from the American Heart Association. *129*, 399-410.

Gong, S., Zheng, C., Doughty, M.L., Losos, K., Didkovsky, N., Schambra, U.B., Nowak, N.J., Joyner, A., Leblanc, G., Hatten, M.E.*, et al.* (2003). A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. Nature *425*, 917-925.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017-1018.

Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847-2849.

Guo, J., Wang, Y., Sachs, F., and Meng, F. (2014). Actin stress in cell reprogramming. Proc Natl Acad Sci U S A *111*, E5252-5261.

Haralick, R.M., Shanmugam, K.J.I.T.o.s., man,, and cybernetics (1973). Textural features for image classification. 610-621.

Haralick, R.M., and Shapiro, L.G. (1992). Computer and robot vision (Addison-wesley).

Hardy, B.S., Uechi, K., Zhen, J., and Pirouz Kavehpour, H. (2009). The deformation of flexible PDMS microchannels under a pressure driven flow. Lab Chip *9*, 935-938.

Heins, N., Malatesta, P., Cecconi, F., Nakafuku, M., Tucker, K.L., Hack, M.A., Chapouton, P., Barde, Y.A., and Gotz, M. (2002). Glial cells generate neurons: the role of the transcription factor Pax6. Nat Neurosci *5*, 308-315.

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.

Heisenberg, C.P., and Bellaiche, Y. (2013). Forces in tissue morphogenesis and patterning. Cell *153*, 948-962.

Hirai, H., Katoku-Kikyo, N., Keirstead, S.A., and Kikyo, N.J.C.r. (2013). Accelerated direct reprogramming of fibroblasts into cardiomyocyte-like cells with the MyoD transactivation domain. *100*, 105-113.

Hirai, H., and Kikyo, N.J.C.r. (2014). Inhibitors of suppressive histone modification promote direct reprogramming of fibroblasts to cardiomyocyte-like cells. *102*, 188-190.

Hollenberg, S.M., Sternglanz, R., Cheng, P.F., and Weintraub, H. (1995). Identification of a new family of tissue-specific basic helix-loop-helix proteins with a two-hybrid system. Mol Cell Biol *15*, 3813-3822.

Hoogaars, W.M., Engel, A., Brons, J.F., Verkerk, A.O., de Lange, F.J., Wong, L.E., Bakker, M.L., Clout, D.E., Wakker, V., Barnett, P.J.G.*, et al.* (2007a). Tbx3 controls the sinoatrial node gene program and imposes pacemaker function on the atria. *21*, 1098-1112.

Hoogaars, W.M., Engel, A., Brons, J.F., Verkerk, A.O., de Lange, F.J., Wong, L.Y., Bakker, M.L., Clout, D.E., Wakker, V., Barnett, P.*, et al.* (2007b). Tbx3 controls

the sinoatrial node gene program and imposes pacemaker function on the atria. Genes Dev *21*, 1098-1112.

Hopcroft, M.A., Nix, W.D., and Kenny, T.W. (2010). What is the Young's Modulus of Silicon? Journal of Microelectromechanical Systems *19*, 229-238.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols *4*, 44.

Ichim, C.V., and Wells, R.A. (2011). Generation of high-titer viral preparations by concentration using successive rounds of ultracentrifugation. J Transl Med *9*, 137.

Ieda, M., Fu, J.-D., Delgado-Olguin, P., Vedantham, V., Hayashi, Y., Bruneau, B.G., and Srivastava, D.J.C. (2010a). Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *142*, 375-386.

Ieda, M., Fu, J.D., Delgado-Olguin, P., Vedantham, V., Hayashi, Y., Bruneau, B.G., and Srivastava, D. (2010b). Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. Cell *142*, 375-386.

Ifkovits, J.L., Addis, R.C., Epstein, J.A., and Gearhart, J.D. (2014a). Inhibition of TGFbeta signaling increases direct conversion of fibroblasts to induced cardiomyocytes. PLoS One *9*, e89678.

Ifkovits, J.L., Addis, R.C., Epstein, J.A., and Gearhart, J.D.J.P.o. (2014b). Inhibition of TGFβ signaling increases direct conversion of fibroblasts to induced cardiomyocytes. *9*, e89678.

Inagawa, K., Miyamoto, K., Yamakawa, H., Muraoka, N., Sadahiro, T., Umei, T., Wada, R., Katsumata, Y., Kaneda, R., and Nakade, K.J.C.r. (2012). Induction of Cardiomyocyte-Like Cells in Infarct Hearts by Gene Transfer of Gata4, Mef2c, and Tbx5Novelty and Significance. *111*, 1147-1156.

Inglis, D.W. (2010). A method for reducing pressure-induced deformation in silicone microfluidics. Biomicrofluidics *4*, 026504.

Jayawardena, T.M., Egemnazarov, B., Finch, E.A., Zhang, L., Payne, J.A., Pandya, K., Zhang, Z., Rosenberg, P., Mirotsou, M., and Dzau, V.J. (2012). MicroRNA-mediated in vitro and in vivo direct reprogramming of cardiac fibroblasts to cardiomyocytes. Circulation research *110*, 1465-1473.

Jones, S.J.G.b. (2004). An overview of the basic helix-loop-helix proteins. *5*, 226.

Jozefczuk, J., Drews, K., and Adjaye, J. (2012). Preparation of mouse embryonic fibroblast cells suitable for culturing human embryonic and induced pluripotent stem cells. J Vis Exp.

Kabsch, W., Mannherz, H.G., Suck, D., Pai, E.F., and Holmes, K.C.J.N. (1990). Atomic structure of the actin: DNase I complex. *347*, 37.

Khan, A., and Mathelier, A. (2017). Intervene: a tool for intersection and visualization of multiple gene or genomic region sets. BMC bioinformatics *18*, 287.

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

Kim, M.M., Huang, Y., Choi, K., and Hidrovo, C.H. (2014). The improved resistance of PDMS to pressure-induced deformation and chemical solvent swelling for microfluidic devices. Microelectronic Engineering *124*, 66-75.

Kotterman, M.A., and Schaffer, D.V. (2014). Engineering adeno-associated viruses for clinical gene therapy. Nat Rev Genet *15*, 445-451.

Koubassova, N., and Tsaturyan, A.J.B. (2011). Molecular mechanism of actin-myosin motor in muscle. *76*, 1484-1506.

Kuo, P.-L., Lee, H., Bray, M.-A., Geisse, N.A., Huang, Y.-T., Adams, W.J., Sheehy, S.P., and Parker, K.K.J.T.A.j.o.p. (2012). Myocyte shape regulates lateral registry of sarcomeres and contractility. *181*, 2030-2037.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods *9*, 357.

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

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

Lee, J.E.J.C.o.i.n. (1997). Basic helix-loop-helix genes in neural development. *7*, 13-20.

Lee, K., Yu, P., Lingampalli, N., Kim, H.J., Tang, R., and Murthy, N.J.I.j.o.n. (2015). Peptide-enhanced mRNA transfection in cultured mouse cardiac fibroblasts and direct reprogramming towards cardiomyocyte-like cells. *10*, 1841.

Li, D., Liu, J., Yang, X., Zhou, C., Guo, J., Wu, C., Qin, Y., Guo, L., He, J., Yu, S.*, et al.* (2017). Chromatin Accessibility Dynamics during iPSC Reprogramming. Cell Stem Cell *21*, 819-833 e816.

Li, X.-H., Li, Q., Jiang, L., Deng, C., Liu, Z., Fu, Y., Zhang, M., Tan, H., Feng, Y., and Shan, Z.J.S.c.t.m. (2015). Generation of Functional Human Cardiac Progenitor Cells by High-Efficiency Protein Transduction. *4*, 1415-1424.

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

Liang, X., Wang, G., Lin, L., Lowe, J., Zhang, Q., Bu, L., Chen, Y., Chen, J., Sun, Y., and Evans, S.M.J.C.r. (2013b). HCN4 Dynamically Marks the First Heart Field and Conduction System PrecursorsNovelty and Significance. *113*, 399-407.

Liang, X., Zhang, Q., Cattaneo, P., Zhuang, S., Gong, X., Spann, N.J., Jiang, C., Cao, X., Zhao, X., Zhang, X.*, et al.* (2015a). Transcription factor ISL1 is essential for pacemaker development and function. J Clin Invest *125*, 3256-3268.

Liang, X., Zhang, Q., Cattaneo, P., Zhuang, S., Gong, X., Spann, N.J., Jiang, C., Cao, X., Zhao, X., and Zhang, X.J.T.J.o.c.i. (2015b). Transcription factor ISL1 is essential for pacemaker development and function. *125*, 3256-3268.

Liao, Y., Smyth, G.K., and Shi, W. (2013). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.

Lin, C.-J., Lin, C.-Y., Chen, C.-H., Zhou, B., and Chang, C.-P.J.D. (2012). Partitioning the heart: mechanisms of cardiac septation and valve development. *139*, 3277-3299.

Lin, Z., and Pu, W.T. (2014). Strategies for cardiac regeneration and repair. Sci Transl Med *6*, 239rv231.

Liu, C., Miller, H., Hui, K.L., Grooman, B., Bolland, S., Upadhyaya, A., and Song, W. (2011). A balance of Bruton's tyrosine kinase and SHIP activation regulates B cell receptor cluster formation by controlling actin remodeling. J Immunol *187*, 230-239.

Liu, N., Barbosa, A.C., Chapman, S.L., Bezprozvannaya, S., Qi, X., Richardson, J.A., Yanagisawa, H., and Olson, E.N.J.D. (2009). DNA binding-dependent and-independent functions of the Hand2 transcription factor during mouse embryogenesis. *136*, 933-942.

Liu, Z., Wang, L., Welch, J.D., Ma, H., Zhou, Y., Vaseghi, H.R., Yu, S., Wall, J.B., Alimohamadi, S., Zheng, M.*, et al.* (2017a). Single-cell transcriptomics reconstructs fate conversion from fibroblast to cardiomyocyte. Nature *551*, 100-104.

Liu, Z., Wang, L., Welch, J.D., Ma, H., Zhou, Y., Vaseghi, H.R., Yu, S., Wall, J.B., Alimohamadi, S., and Zheng, M.J.N. (2017b). Single-cell transcriptomics reconstructs fate conversion from fibroblast to cardiomyocyte. *551*, 100.

Long-Fang Tsai, W.C.D., Seunghyun Kim, Gregory P. Nordin (2011). Bonding of polydimethylsiloxane microfluidics to silicon-based sensors. Journal of Micro/Nanolithography, MEMS, and MOEMS *1439*.

Long, H.K., Prescott, S.L., and Wysocka, J. (2016). Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. Cell *167*, 1170-1187.

Luni, C., Giulitti, S., Serena, E., Ferrari, L., Zambon, A., Gagliano, O., Giobbe, G.G., Michielin, F., Knobel, S., Bosio, A.*, et al.* (2016). High-efficiency cellular reprogramming with microfluidics. Nat Methods *13*, 446-452.

Ma, Y., Iyer, R.P., Jung, M., Czubryt, M.P., and Lindsey, M.L.J.T.i.p.s. (2017). Cardiac fibroblast activation post-myocardial infarction: current knowledge gaps. *38*, 448-458.

Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R.*, et al.* (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat Neurosci *13*, 133-140.

Mall, M., Kareta, M.S., Chanda, S., Ahlenius, H., Perotti, N., Zhou, B., Grieder, S.D., Ge, X., Drake, S., Euong Ang, C.*, et al.* (2017). Myt1l safeguards neuronal identity by actively repressing many non-neuronal fates. Nature *544*, 245-249.

Martin-Puig, S., Wang, Z., and Chien, K.R.J.C.s.c. (2008). Lives of a heart cell: tracing the origins of cardiac progenitors. *2*, 320-331.

McBeath, R., Pirone, D.M., Nelson, C.M., Bhadriraju, K., and Chen, C.S. (2004). Cell shape, cytoskeletal tension, and RhoA regulate stem cell lineage commitment. Dev Cell *6*, 483-495.

McDonald, J.C., and Whitesides, G.M. (2002). Poly(dimethylsiloxane) as a material for fabricating microfluidic devices. Acc Chem Res *35*, 491-499.

McFadden, D.G., McAnally, J., Richardson, J.A., Charite, J., and Olson, E.N. (2002). Misexpression of dHAND induces ectopic digits in the developing limb bud in the absence of direct DNA binding. Development *129*, 3077-3088.

McGough, A., Way, M., and DeRosier, D.J.T.J.o.C.B. (1994). Determination of the alpha-actinin-binding site on actin filaments by cryoelectron microscopy and image analysis. *126*, 433-443.

Meng, F., Suchyna, T.M., Lazakovitch, E., Gronostajski, R.M., and Sachs, F. (2011). Real Time FRET Based Detection of Mechanical Stress in Cytoskeletal and Extracellular Matrix Proteins. Cell Mol Bioeng *4*, 148-159.

Meng, F., Suchyna, T.M., and Sachs, F. (2008). A fluorescence energy transfer-based mechanical stress sensor for specific proteins in situ. FEBS J *275*, 3072-3087.

Merchant, N., Lyons, E., Goff, S., Vaughn, M., Ware, D., Micklos, D., and Antin, P.J.P.b. (2016). The iPlant collaborative: cyberinfrastructure for enabling data to discovery for the life sciences. *14*, e1002342.

Mezzano, V., Liang, Y., Wright, A.T., Lyon, R.C., Pfeiffer, E., Song, M.Y., Gu, Y., Dalton, N.D., Scheinman, M., Peterson, K.L.*, et al.* (2016). Desmosomal junctions are necessary for adult sinus node function. Cardiovasc Res *111*, 274-286.

Miki, K., Yoshida, Y., and Yamanaka, S. (2013). Making steady progress on direct cardiac reprogramming toward clinical application. Circulation research *113*, 13-15.

Miquerol, L., Beyer, S., and Kelly, R.G. (2011). Establishment of the mouse ventricular conduction system. Cardiovasc Res *91*, 232-242.

Miyamoto, K., Akiyama, M., Tamura, F., Isomi, M., Yamakawa, H., Sadahiro, T., Muraoka, N., Kojima, H., Haginiwa, S., Kurotsu, S.*, et al.* (2018). Direct In Vivo Reprogramming with Sendai Virus Vectors Improves Cardiac Function after Myocardial Infarction. Cell Stem Cell *22*, 91-103 e105.

Mohamed, T.M., Stone, N.R., Berry, E.C., Radzinsky, E., Huang, Y., Pratt, K., Ang, Y.S., Yu, P., Wang, H., Tang, S.*, et al.* (2017). Chemical Enhancement of In Vitro and In Vivo Direct Cardiac Reprogramming. Circulation *135*, 978-995.

Mommersteeg, M.T., Hoogaars, W.M., Prall, O.W., de Gier-de Vries, C., Wiese, C., Clout, D.E., Papaioannou, V.E., Brown, N.A., Harvey, R.P., Moorman, A.F.*, et al.* (2007a). Molecular pathway for the localized formation of the sinoatrial node. Circ Res *100*, 354-362.

Mommersteeg, M.T., Hoogaars, W.M., Prall, O.W., de Gier-de Vries, C., Wiese, C., Clout, D.E., Papaioannou, V.E., Brown, N.A., Harvey, R.P., and Moorman, A.F.J.C.r. (2007b). Molecular pathway for the localized formation of the sinoatrial node. *100*, 354-362.

Morin, S., Charron, F., Robitaille, L., and Nemer, M. (2000). GATA-dependent recruitment of MEF2 proteins to target promoters. EMBO J *19*, 2046-2055.

Morris, S.A. (2016). Direct lineage reprogramming via pioneer factors; a detour through developmental gene regulatory networks. Development *143*, 2696-2705.

Muraoka, N., and Ieda, M. (2014). Direct reprogramming of fibroblasts into myocytes to reverse fibrosis. Annu Rev Physiol *76*, 21-37.

Muraoka, N., Yamakawa, H., Miyamoto, K., Sadahiro, T., Umei, T., Isomi, M., Nakashima, H., Akiyama, M., Wada, R., Inagawa, K.*, et al.* (2014a). MiR-133

promotes cardiac reprogramming by directly repressing Snai1 and silencing fibroblast signatures. EMBO J *33*, 1565-1581.

Muraoka, N., Yamakawa, H., Miyamoto, K., Sadahiro, T., Umei, T., Isomi, M., Nakashima, H., Akiyama, M., Wada, R., and Inagawa, K.J.T.E.j. (2014b). MiR-133 promotes cardiac reprogramming by directly repressing Snai1 and silencing fibroblast signatures. *33*, 1565-1581.

Murre, C., Bain, G., van Dijk, M.A., Engel, I., Furnari, B.A., Massari, M.E., Matthews, J.R., Quong, M.W., Rivera, R.R., Stuiver, M.H.J.B.e.B.A.-G.S*., et al.* (1994). Structure and function of helix-loop-helix proteins. *1218*, 129-135.

Murre, C., McCaw, P.S., Vaessin, H., Caudy, M., Jan, L., Jan, Y., Cabrera, C.V., Buskin, J.N., Hauschka, S.D., and Lassar, A.B.J.C. (1989). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *58*, 537-544.

Nam, Y.-J., Lubczyk, C., Bhakta, M., Zang, T., Fernandez-Perez, A., McAnally, J., Bassel-Duby, R., Olson, E.N., and Munshi, N.V.J.D. (2014a). Induction of diverse cardiac cell types by reprogramming fibroblasts with cardiac transcription factors. dev. 114025.

Nam, Y.J., Lubczyk, C., Bhakta, M., Zang, T., Fernandez-Perez, A., McAnally, J., Bassel-Duby, R., Olson, E.N., and Munshi, N.V. (2014b). Induction of diverse cardiac cell types by reprogramming fibroblasts with cardiac transcription factors. Development *141*, 4267-4278.

Nam, Y.J., Song, K., Luo, X., Daniel, E., Lambeth, K., West, K., Hill, J.A., DiMaio, J.M., Baker, L.A., Bassel-Duby, R*., et al.* (2013). Reprogramming of human fibroblasts toward a cardiac fate. Proc Natl Acad Sci U S A *110*, 5588-5593.

Otsu, N.J.I.t.o.s., man,, and cybernetics (1979). A threshold selection method from gray-level histograms. *9*, 62-66.

Pappas, C.T., Bhattacharya, N., Cooper, J.A., and Gregorio, C.C.J.M.b.o.t.c. (2008). Nebulin interacts with CapZ and regulates thin filament architecture within the Z-disc. *19*, 1837-1847.

Pasqualini, F.S., Sheehy, S.P., Agarwal, A., Aratyn-Schaus, Y., and Parker, K.K.J.S.c.r. (2015). Structural phenotyping of stem cell-derived cardiomyocytes. *4*, 340-347.

Pereira, C.F., Chang, B., Gomes, A., Bernitz, J., Papatsenko, D., Niu, X., Swiers, G., Azzoni, E., de Bruijn, M.F., Schaniel, C*., et al.* (2016). Hematopoietic Reprogramming In Vitro Informs In Vivo Identification of Hemogenic Precursors to Definitive Hematopoietic Stem Cells. Dev Cell *36*, 525-539.

Prabhu, S.D., and Frangogiannis, N.G.J.C.r. (2016). The biological basis for cardiac repair after myocardial infarction: from inflammation to fibrosis. *119*, 91-112.

Qian, L., Berry, E.C., Fu, J.D., Ieda, M., and Srivastava, D. (2013). Reprogramming of mouse fibroblasts into cardiomyocyte-like cells in vitro. Nat Protoc *8*, 1204-1215.

Qian, L., Huang, Y., Spencer, C.I., Foley, A., Vedantham, V., Liu, L., Conway, S.J., Fu, J.-d., and Srivastava, D.J.N. (2012a). In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. *485*, 593.

Qian, L., Huang, Y., Spencer, C.I., Foley, A., Vedantham, V., Liu, L., Conway, S.J., Fu, J.D., and Srivastava, D. (2012b). In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. Nature *485*, 593-598.

Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic acids research *44*, W160-165.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome biology *11*, R25.

Sadahiro, T., Yamanaka, S., and Ieda, M.J.C.r. (2015). Direct cardiac reprogramming: progress and challenges in basic biology and clinical applications. *116*, 1378-1391.

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.

Sharei, A., Zoldan, J., Adamo, A., Sim, W.Y., Cho, N., Jackson, E., Mao, S., Schneider, S., Han, M.J., Lytton-Jean, A.*, et al.* (2013). A vector-free microfluidic platform for intracellular delivery. Proc Natl Acad Sci U S A *110*, 2082-2087.

Sheehy, S.P., Pasqualini, F., Grosberg, A., Park, S.J., Aratyn-Schaus, Y., and Parker, K.K.J.S.c.r. (2014). Quality metrics for stem cell-derived cardiac myocytes. *2*, 282-294.

Shiba, Y., Fernandes, S., Zhu, W.-Z., Filice, D., Muskheli, V., Kim, J., Palpant, N.J., Gantz, J., Moyes, K.W., and Reinecke, H.J.N. (2012). Human ES-cell-derived cardiomyocytes electrically couple and suppress arrhythmias in injured hearts. *489*, 322.

Shimp, E.A., Alsmadi, N.Z., Cheng, T., Lam, K.H., Lewis, C.S., and Schmidtke, D.W. (2016). Effects of shear on P-selectin deposition in microfluidic channels. Biomicrofluidics *10*, 024128.

Shiroma, L.S., Piazzetta, M.H., Duarte-Junior, G.F., Coltro, W.K., Carrilho, E., Gobbi, A.L., and Lima, R.S. (2016). Self-regenerating and hybrid irreversible/reversible PDMS microfluidic devices. Sci Rep *6*, 26032.

Shubeita, H.E., McDonough, P.M., Harris, A., Knowlton, K., Glembotski, C., Brown, J.H., and Chien, K.J.J.o.B.C. (1990). Endothelin induction of inositol phospholipid hydrolysis, sarcomere assembly, and cardiac gene expression in ventricular myocytes. A paracrine mechanism for myocardial cell hypertrophy. *265*, 20555-20562.

Sissman, N.J.J.A.J.o.C. (1970). Developmental landmarks in cardiac morphogenesis: comparative chronology. *25*, 141-148.

Sjöblom, B., Salmazo, A., Djinović-Carugo, K.J.C., and sciences, m.l. (2008). α-Actinin structure and regulation. *65*, 2688.

Smith, Z.D., Sindhu, C., and Meissner, A. (2016). Molecular features of cellular reprogramming and development. Nature reviews Molecular cell biology *17*, 139-154.

Song, K., Nam, Y.-J., Luo, X., Qi, X., Tan, W., Huang, G.N., Acharya, A., Smith, C.L., Tallquist, M.D., and Neilson, E.G.J.N. (2012a). Heart repair by reprogramming non-myocytes with cardiac transcription factors. *485*, 599.

Song, K., Nam, Y.J., Luo, X., Qi, X., Tan, W., Huang, G.N., Acharya, A., Smith, C.L., Tallquist, M.D., Neilson, E.G.*, et al.* (2012b). Heart repair by reprogramming non-myocytes with cardiac transcription factors. Nature *485*, 599-604.

Song, K.H., Nam, Y.J., Luo, X., Qi, X.X., Tan, W., Huang, G.N., Acharya, A., Smith, C.L., Tallquist, M.D., Neilson, E.G.*, et al.* (2012c). Heart repair by reprogramming non-myocytes with cardiac transcription factors. Nature *485*, 599-+.

Sparrow, J.C., and Schöck, F.J.N.R.M.C.B. (2009). The initial steps of myofibril assembly: integrins pave the way. *10*, 293.

Spater, D., Hansson, E.M., Zangi, L., and Chien, K.R. (2014). How to make a cardiomyocyte. Development *141*, 4418-4431.

Sreejit, P., Kumar, S., Verma, R.S.J.I.V.C., and Biology-Animal, D. (2008). An improved protocol for primary culture of cardiomyocyte from neonatal mice. *44*, 45-50.

Srivastava, D., Cserjesi, P., and Olson, E.N. (1995). A subclass of bHLH proteins required for cardiac morphogenesis. Science *270*, 1995-1999.

Srivastava, D., and DeWitt, N. (2016a). In Vivo Cellular Reprogramming: The Next Generation. Cell *166*, 1386-1396.

Srivastava, D., and DeWitt, N.J.C. (2016b). In vivo cellular reprogramming: the next generation. *166*, 1386-1396.

Srivastava, D., and Olson, E.N.J.T.i.c.b. (1997). Knowing in your heart what's right. *7*, 447-453.

Srivastava, D., Thomas, T., Lin, Q., Kirby, M.L., Brown, D., and Olson, E.N. (1997). Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND. Nat Genet *16*, 154-160.

Stankunas, K., Shang, C., Twu, K.Y., Kao, S.C., Jenkins, N.A., Copeland, N.G., Sanyal, M., Selleri, L., Cleary, M.L., and Chang, C.P. (2008). Pbx/Meis deficiencies demonstrate multigenetic origins of congenital heart disease. Circ Res *103*, 702-709.

Stanley, E.G., Biben, C., Elefanty, A., Barnett, L., Koentgen, F., Robb, L., and Harvey, R.P.J.I.J.o.D.B. (2004). Efficient Cre-mediated deletion in cardiac progenitor cells conferred by a 3'UTR-ires-Cre allele of the homeobox gene Nkx2-5. *46*, 431-439.

Stark, R., and Brown, G. (2011). DiffBind: differential binding analysis of ChIP-Seq peak data. R package version *100*, 4.3.

Stewart, M.P., Sharei, A., Ding, X., Sahay, G., Langer, R., and Jensen, K.F. (2016). In vitro and ex vivo strategies for intracellular delivery. Nature *538*, 183-192.

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.

Talkhabi, M., Pahlavan, S., Aghdami, N., Baharvand, H.J.B., and communications, b.r. (2015). Ascorbic acid promotes the direct conversion of mouse fibroblasts into beating cardiomyocytes. *463*, 699-705.

Togi, K., Yoshida, Y., Matsumae, H., Nakashima, Y., Kita, T., Tanaka, M.J.B., and communications, b.r. (2006). Essential role of Hand2 in interventricular septum formation and trabeculation during cardiac development. *343*, 144-151.

Treutlein, B., Lee, Q.Y., Camp, J.G., Mall, M., Koh, W., Shariati, S.A., Sim, S., Neff, N.F., Skotheim, J.M., Wernig, M.*, et al.* (2016). Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. Nature *534*, 391-395.

Tyser, R.C., Miranda, A.M., Chen, C.-m., Davidson, S.M., Srinivas, S., and Riley, P.R.J.E. (2016). Calcium handling precedes cardiac differentiation to initiate the first heartbeat. *5*, e17113.

Tzahor, E., and Poss, K.D.J.S. (2017). Cardiac regeneration strategies: staying young at heart. *356*, 1035-1039.

van Eif, V.W., Devalla, H.D., Boink, G.J., and Christoffels, V.M.J.N.R.C. (2018a). Transcriptional regulation of the cardiac conduction system. 1.

van Eif, V.W.W., Devalla, H.D., Boink, G.J.J., and Christoffels, V.M. (2018b). Transcriptional regulation of the cardiac conduction system. Nature reviews Cardiology *15*, 617-630.

VanDusen, N.J., Vincentz, J.W., Firulli, B.A., Howard, M.J., Rubart, M., and Firulli, A.B.J.D.b. (2014). Loss of Hand2 in a population of Periostin lineage cells results in pronounced bradycardia and neonatal death. *388*, 149-158.

Vavylonis, D., Yang, Q., and O'Shaughnessy, B.J.P.o.t.N.A.o.S. (2005). Actin polymerization kinetics, cap structure, and fluctuations. *102*, 8543-8548.

Vedantham, V., Galang, G., Evangelista, M., Deo, R.C., and Srivastava, D. (2015). RNA sequencing of mouse sinoatrial node reveals an upstream regulatory role for Islet-1 in cardiac pacemaker cells. Circ Res *116*, 797-803.

Verdurmen, W.P.R., Mazlami, M., and Pluckthun, A. (2017). A quantitative comparison of cytosolic delivery via different protein uptake systems. Scientific reports *7*, 13194.

Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Sudhof, T.C., and Wernig, M. (2010a). Direct conversion of fibroblasts to functional neurons by defined factors. Nature *463*, 1035-U1050.

Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Südhof, T.C., and Wernig, M. (2010b). Direct conversion of fibroblasts to functional neurons by defined factors. Nature *463*, 1035-1041.

Vincent, L., Soille, P.J.I.T.o.P.A., and Intelligence, M. (1991). Watersheds in digital spaces: an efficient algorithm based on immersion simulations. 583-598.

Vincentz, J.W., Barnes, R.M., Firulli, A.B.J.B.D.R.P.A.C., and Teratology, M. (2011). Hand factors as regulators of cardiac morphogenesis and implications for congenital heart defects. *91*, 485-494.

Wada, R., Muraoka, N., Inagawa, K., Yamakawa, H., Miyamoto, K., Sadahiro, T., Umei, T., Kaneda, R., Suzuki, T., and Kamiya, K.J.P.o.t.N.A.o.S. (2013). Induction of human cardiomyocyte-like cells from fibroblasts by defined factors. *110*, 12667-12672.

Wang, H., Cao, N., Spencer, C.I., Nie, B., Ma, T., Xu, T., Zhang, Y., Wang, X., Srivastava, D., and Ding, S.J.C.r. (2014a). Small molecules enable cardiac reprogramming of mouse fibroblasts with a single factor, Oct4. *6*, 951-960.

Wang, J., Chen, W., Sun, J., Liu, C., Yin, Q., Zhang, L., Xianyu, Y., Shi, X., Hu, G., and Jiang, X. (2014b). A microfluidic tubing method and its application for controlled synthesis of polymeric nanoparticles. Lab on a Chip *14*, 1673-1677.

Wang, L., Liu, Z., Yin, C., Asfour, H., Chen, O., Li, Y., Bursac, N., Liu, J., and Qian, L.J.C.r. (2015). Stoichiometry of Gata4, Mef2c, and Tbx5 influences the efficiency and quality of induced cardiac myocyte reprogramming. *116*, 237-244.

Wang, L.H., and Baker, N.E. (2015). E Proteins and ID Proteins: Helix-Loop-Helix Partners in Development and Disease. Dev Cell *35*, 269-280.

Wapinski, O.L., Lee, Q.Y., Chen, A.C., Li, R., Corces, M.R., Ang, C.E., Treutlein, B., Xiang, C., Baubet, V., Suchy, F.P.*, et al.* (2017). Rapid Chromatin Switch in the Direct Reprogramming of Fibroblasts to Neurons. Cell reports *20*, 3236-3247.

Wapinski, O.L., Vierbuchen, T., Qu, K., Lee, Q.Y., Chanda, S., Fuentes, D.R., Giresi, P.G., Ng, Y.H., Marro, S., Neff, N.F.*, et al.* (2013). Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. Cell *155*, 621-635.

Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Liaw, W.H.A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., and Schwartz, M. (2016). gplots: various R programming tools for plotting data. R package version 3.0. 1. The Comprehensive R Archive Network.

Weintraub, H., Dwarki, V., Verma, I., Davis, R., Hollenberg, S., Snider, L., Lassar, A., Tapscott, S.J.J.G., and development (1991). Muscle-specific transcriptional activation by MyoD. *5*, 1377-1386.

Whitesides, G.M., and Tang, S.K.Y. (2006). Fluidic optics.

Wiese, C., Grieskamp, T., Airik, R., Mommersteeg, M.T., Gardiwal, A., de Gier-de Vries, C., Schuster-Gossler, K., Moorman, A.F., Kispert, A., and Christoffels, V.M.J.C.r. (2009). Formation of the sinus node head and differentiation of sinus node myocardium are independently regulated by Tbx18 and Tbx3. *104*, 388-397.

Wingett, S. (2011). FastQ Screen: A quality control tool to screen a library of sequences in FastQ format against a set of sequence databases.

Writing Group, M., Mozaffarian, D., Benjamin, E.J., Go, A.S., Arnett, D.K., Blaha, M.J., Cushman, M., Das, S.R., de Ferranti, S., Despres, J.P.*, et al.* (2016). Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association. Circulation *133*, e38-360.

Xin, M., Olson, E.N., and Bassel-Duby, R.J.N.r.M.c.b. (2013). Mending broken hearts: cardiac development as a basis for adult heart regeneration and repair. *14*, 529.

Yamada, M., Kano, K., Tsuda, Y., Kobayashi, J., Yamato, M., Seki, M., and Okano, T. (2007). Microfluidic devices for size-dependent separation of liver cells. Biomed Microdevices *9*, 637-645.

Yamakawa, H., Muraoka, N., Miyamoto, K., Sadahiro, T., Isomi, M., Haginiwa, S., Kojima, H., Umei, T., Akiyama, M., Kuishi, Y.*, et al.* (2015a). Fibroblast Growth

Factors and Vascular Endothelial Growth Factor Promote Cardiac Reprogramming under Defined Conditions. Stem cell reports *5*, 1128-1142.

Yamakawa, H., Muraoka, N., Miyamoto, K., Sadahiro, T., Isomi, M., Haginiwa, S., Kojima, H., Umei, T., Akiyama, M., and Kuishi, Y.J.S.c.r. (2015b). Fibroblast growth factors and vascular endothelial growth factor promote cardiac reprogramming under defined conditions. *5*, 1128-1142.

Yang, R., Zheng, Y., Li, L., Liu, S., Burrows, M., Wei, Z., Nace, A., Herlyn, M., Cui, R., Guo, W.*, et al.* (2014a). Direct conversion of mouse and human fibroblasts to functional melanocytes by defined factors. Nat Commun *5*, 5807.

Yang, X., Pabon, L., and Murry, C.E.J.C.r. (2014b). Engineering adolescence: maturation of human pluripotent stem cell–derived cardiomyocytes. *114*, 511-523.

Yoshida, Y., and Yamanaka, S.J.C.r. (2017). Induced pluripotent stem cells 10 years later: for cardiac applications. *120*, 1958-1968.

Young, P., Ferguson, C., Bañuelos, S., and Gautel, M.J.T.E.j. (1998). Molecular structure of the sarcomeric Z-disk: two types of titin interactions lead to an asymmetrical sorting of α-actinin. *17*, 1614-1624.

Yu, G., Wang, L.G., and He, Q.Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics *31*, 2382-2383.

Zang, M.X., Li, Y., Wang, H., Wang, J.B., and Jia, H.T. (2004). Cooperative interaction between the basic helix-loop-helix transcription factor dHAND and myocyte enhancer factor 2C regulates myocardial gene expression. J Biol Chem *279*, 54258-54263.

Zebrowski, D.C., Vergarajauregui, S., Wu, C.-C., Piatkowski, T., Becker, R., Leone, M., Hirth, S., Ricciardi, F., Falk, N., and Giessl, A.J.E. (2015). Developmental alterations in centrosome integrity contribute to the post-mitotic state of mammalian cardiomyocytes. *4*, e05563.

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

Zhang, Y., Zhong, J.F., Qiu, H., MacLellan, W.R., Marbán, E., and Wang, C.J.S.r. (2015). Epigenomic reprogramming of adult cardiomyocyte-derived cardiac progenitor cells. *5*, 17686.

Zhao, Y., Londono, P., Cao, Y., Sharpe, E.J., Proenza, C., O'Rourke, R., Jones, K.L., Jeong, M.Y., Walker, L.A., Buttrick, P.M.*, et al.* (2015). High-efficiency

reprogramming of fibroblasts into cardiomyocytes requires suppression of pro-fibrotic signalling. Nat Commun *6*, 8243.

Zhou, H., Dickson, M.E., Kim, M.S., Bassel-Duby, R., and Olson, E.N. (2015). Akt1/protein kinase B enhances transcriptional reprogramming of fibroblasts to functional cardiomyocytes. Proc Natl Acad Sci U S A *112*, 11864-11869.

Zhou, H., Morales, M.G., Hashimoto, H., Dickson, M.E., Song, K., Ye, W., Kim, M.S., Niederstrasser, H., Wang, Z., Chen, B.*, et al.* (2017). ZNF281 enhances cardiac reprogramming by modulating cardiac and inflammatory gene expression. Genes Dev *31*, 1770-1783.

Zhou, Y., Wang, L., Vaseghi, H.R., Liu, Z., Lu, R., Alimohamadi, S., Yin, C., Fu, J.D., Wang, G.G., Liu, J.*, et al.* (2016). Bmi1 Is a Key Epigenetic Barrier to Direct Cardiac Reprogramming. Cell Stem Cell *18*, 382-395.

Zúñiga, R., González, D., Valenzuela, C., Brown, N., Zúñiga, L.J.B., and communications, b.r. (2016). Expression and cellular localization of HCN channels in rat cerebellar granule neurons. *478*, 1429-1435.