Title
Harnessing the Power of Human Embryonic Stem Cells and Direct Reprogramming for Cardiac Regeneration

Author
Engel, James

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Harnessing the Power of Human Embryonic Stem Cells and Direct Reprogramming for Cardiac Regeneration

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Molecular, Cellular and Integrative Physiology

by

James Lloyd Engel Jr.

2019

ABSTRACT OF THE DISSERTATION

Harnessing the Power of Human Embryonic Stem Cells and Direct Reprogramming for Cardiac Regeneration

by

James Lloyd Engel Jr.

Doctor of Philosophy in Molecular, Cellular and Integrative Physiology

University of California, Los Angeles, 2019

Professor Reza Ardehali, Chair

Acquiring pure populations of cardiomyocyte (CM) subtypes is important for developmental studies and necessary for safe cell-based therapies. Here we generated cardiac-specific human embryonic stem cell (hESC) reporter lines to isolate first and second heart field-like CMs (FHF and SHF-like respectively) and cells of the conduction system. Our studies demonstrate that FHF-like CMs can be prospectively isolated based on the co-expression of TBX5/NKX2-5 whereas SHF-like CMs are marked by NKX2-5 but lack TBX5 expression, and finally nodal cells are enriched within the fraction expressing only TBX5. Characterization of these cell populations by electrophysiological, functional, transplantation, and transcriptional studies supports the cellular identity assigned on the basis of NKX2-5 and TBX5 expression.

Furthermore, we identify CORIN as a novel cell surface marker of both FHF-like progenitors and CMs. These results provide a platform to investigate *in vitro* cardiovascular development and may facilitate a safe approach for cell therapy in heart disease.

In addition to the use of hESC derived cardiomyocyte-like cells to repopulate the heart and fully restore cardiac function after insult, we explore the use direct reprogramming to convert fibroblasts into cardiomyocyte-like cells. Direct cardiac reprogramming using transcription factors, small molecules, miRNAs, and other biologics for the treatment of heart failure have been explored. Here, we focus on the use of modified RNAs and small molecules to develop a safe and efficient method of direct reprogramming for the treatment of heart failure.

The dissertation of James Lloyd Engel Jr. is approved.

April Dawn Pyle

Yibin Wang

James N. Weiss

Reza Ardehali, Chair

University of California, Los Angeles

2019

Dedication

A successful graduate career only occurs with a group of exceptional family, peers and mentors that provide their unyielding support. First and foremost, I would like to thank my family for all their support. Mom and Dad, you have been there for me throughout all my endeavors and provided support when times were tough. April, you have encouraged me to enjoy life and have provided a great getaway in San Diego.

I thank my supervisor and mentor Dr. Reza Ardehali. Throughout my graduate studies you have been there to support and guide me during both the good times and the bad times. Your enthusiasm for translational research and dedication to improving the lives of those with heart disease is exceptional. I also would like to thank of the members of my committee for their support and guidance.

Last but certainly not least, I thank the past and present members of the Ardehali lab. You guys have been the best colleagues, coworkers, friends and family that anyone could wish for. I cherish the memories and friendships we have developed inside and outside the lab and without you guys this would not be possible.

Table of Contents

Chapter 1:

Isolation of Heart Field Specific Cardiomyocytes from Differentiating Human Embryonic Stem Cells for Cardiac Regeneration

Chapter 2:

Direct Cardiac Reprogramming

List of Figures

List of Tables

Acknowledgments

Chapter 1 is a version of a manuscript currently under revision at Cell Reports.

Engel JL, Pezhouman A, Khoja S, Nguyen NB, Skelton RJP, Gilmore WB, Sahoo D, Elliott DA, Zhao P and Ardehali R. Isolation of Heart Field Specific Cardiomyocytes from Differentiating Human Embryonic Stem Cells for Cardiac Regeneration. Cell Reports. In Revision.

Author Contributions:

Conceptualization, J.L.E., A.P., S.K. and R.A.; Methodology, J.L.E., A.P., S.K., R.J.P.S., W.B.G., and R.A.; Formal Analysis, N.B.N. and D.S.; Investigation. J.L.E, A.P., S.K., R.J.P.S., W.B.G. and P.Z.; Resources, D.S, and D.A.E.; Writing-Original Draft, J.L.E., A.P. and R.A.; Writing- Review & Editing, J.L.E., A.P., N.B.N., D.A.E., P.Z. and R.A.; Visualization, J.L.E., A.P., S.K., N.B.N., P.Z., R.A.; Supervision, R.A.; Funding Acquisition, R.A.

Chapter 2 is a version of a published review by Engel and Ardehali.

Engel JL and Ardehali R. Direct Cardiac Reprogramming: Progress and Promise. Stem Cells Int. 2018 Mar 13;2018:1435746. doi: 10.1155/2018/1435746. eCollection 2018. Review.

Vita

Education

University of California, San Diego, 2007-2009

- M.S. in Biology
- Research Advisor: Dr. Jack Dixon
- Thesis Title: The *Pseudomonas syringae* type III effector HopE1 interacts and inhibits an *Arabidopsis* WNK Kinase

University of California, San Diego, 2004-2007

- B.S. in Human Biology

Riverside Community College, 2002-2004

Publications:

Engel JL, Ardehali R. Sendai virus based direct cardiac reprogramming: what lies ahead? Stem Cell Investig. 2018 Oct 22;5:37.

Engel JL, Ardehali R. Direct Cardiac Reprogramming: Progress and Promise. Stem Cells Int. 2018 Mar 13;2018. Review.

Sereti KI, Nguyen NB, Kamran P, Zhao P, Ranjbarvaziri S, Park S, Sabri S, **Engel JL**, Sung K, Kulkarni RP, Ding Y, Hsiai TK, Plath K, Ernst J, Sahoo D, Mikkola HKA, Iruela-Arispe ML, Ardehali R. Analysis of cardiomyocyte clonal expansion during mouse heart development and injury. Nat Commun. 2018 Feb 21;9(1):754.

Wang X, Cimermancic P, Yu C, Schweitzer A, Chopra N, **Engel JL**, Greenberg C, Huszagh AS, Beck F, Sakata E, Yang Y, Novitsky EJ, Leitner A, Nanni P, Kahraman A, Guo X, Dixon JE, Rychnovsky SD, Aebersold R, Baumeister W, Sali A, Huang L. Molecular Details Underlying Dynamic Structures and Regulation of the Human 26S Proteasome. Mol Cell Proteomics. 2017 May;16(5):840-854.

Skelton RJ, Brady B, Sahoo D, **Engel JL**, Abilez O, Hojjat A, Khoja S, Zhao P, Kwon M, Elliott DA, Ardehali R. CD13 and ROR2 permit isolation of highly enriched cardiac mesoderm from human embryonic stem cells. Stem Cell Reports. 2016 Jan 12;6(1):95-108.

Skelton RJ, Almeida S, Rapacchi S, Han F, Khoja S, **Engel JL**, Zhao P, Peng H, Stanley E, Elefanty A, Kwon M, Elliot DA, Ardehali R. Magnetic resonance imaging of iron oxide labeled human embryonic stem cell derived cardiac progenitors. Stem Cells Transl Med. 2016 Jan;5(1):67-74.

Tagliabracci VS, Wiley SE, Guo X, Kinch LN, Durrant E, Wen J, Xiao J, Cui J, **Engel JL**, Coon JL, Grishin NV, Pinna LA, Pagiliarini DJ, Dixon JE. A single kinase generates the majority of the secreted phosphoproteome. Cell. Jun 18;161(7):1619-32.

Tagliabracci VS, **Engel JL**, Wiley SW, Xiao J, Gonzalez DJ, Nidumanda Appaiah H, Koller A, Nizet V, White KE, Dixon JE. Dynamic regulation of FGF23 by Fam20C phosphorylation, GalNAc-T3 glycosylation, and furin proteolysis. Proc Natl Acad Sci USA. 2014 Apr 15;111(15):5520-5.

Teh PG, Chen MJ, **Engel JL**, Worby CA, Manning G, Dixon JE, Zhang J. Identification of a Mammalian-type Phosphatidylglycerophosphate Phosphatase in the Eubacterium Rhodopirellula baltica. J Biol Chem. 2013 Feb 15;288(7):5176-85.

Tagliabracci VS, **Engel JL**, Wen J, Wiley SW, Worby CA, Kinch LN, Xiao J, Grishin NV, Dixon JE. Secreted kinase phosphorylates extracellular proteins that regulate biomineralization. Science. 2012 Jun 1;336(6085):1150-3

Guo X, **Engel JL**, Xiao J, Tagliabracci VS, Wang X, Huang L, Dixon JE. UBLCP1 is a 26S proteasome phosphatase that regulates nuclear proteasome activity. Proc Natl Acad Sci USA. 2011 Nov 15;108(46):18649-54.

Xiao J*****, **Engel JL***, Zhang J, Chen MJ, Manning G, Dixon JE. Structural and functional analysis of PTPMT1, a phosphatase required for cardiolipin synthesis. Proc Natl Acad Sci U S A. 2011 Jul 19;108(29):11860-5.

*** Authors contributed equally to this work**

Zhang J, Guan Z, Murphy AN, Wiley SE, Perkins GA, Worby CA, **Engel JL**, Heacock P, Nguyen OK, Wang JH, Raetz CR, Dowhan W, Dixon JE. Mitochondrial phosphatase PTPMT1 is essential for cardiolipin biosynthesis. Cell Metab. 2011 Jun 8;13(6):690-700.

Dowen RH, **Engel JL**, Shao F, Ecker JR, Dixon JE. A family of bacterial cysteine protease type III effectors utilizes acylation-dependent and -independent strategies to localize to plasma membranes. J Biol Chem. 2009 Jun 5;284(23):15867-79.

Chapter 1:

Isolation of Heart Field Specific Cardiomyocytes from Differentiating Human Embryonic Stem Cells for Cardiac Regeneration

Abstract

Current hESC differentiation protocols yield a mixture of cardiac cells without the ability to isolate heart field specific cardiac progenitors and their associated progeny. Acquiring pure populations of cardiomyocyte (CM) subtypes is important for developmental studies and necessary for safe cell-based therapies. Here, we generated cardiac-specific human embryonic stem cell (hESC) reporter lines to isolate first and second heart field-like CMs (FHF and SHF-like respectively) and cells of the conduction system. The cell populations were characterized using immunostaining, qPCR and electrophysiological studies.

Using our FHF specific TBX5-$^{TdTomato/W}$ hESC reporter line we show that TBX5^{+} cells represent an enriched population of FHF CPCs that can give rise to CMs, endothelial and smooth muscle cells *in vitro*. Transplantation of hESC-derived TBX5^{+} and TBX5^{-} CPCs into the injured left ventricle of mice showed that these cells were able to survive and predominantly gave rise to CMs. RNA-sequencing analysis of hESC-derived cardiac cells at different stages of development suggested that while TBX5 expression leads to an enrichment of FHF- derived CMs, TBX5^{-} cells also generate cardiac cells that resemble SHF CMs at the transcriptional level. To enable prospective isolation of FHF- and SHF-like CMs we generated a double transgenic hESC reporter line TBX5$^{TdTomato/W}$/NKX2-5$^{eGFP/W}$. Our studies demonstrate that FHF-like CMs can be prospectively isolated based on the co-expression of TBX5/NKX2-5 whereas SHF-like CMs are marked by NKX2-5 but lack TBX5 expression, and finally nodal cells are enriched within the fraction expressing only TBX5. Using electrophysiological, functional, transplantation, and transcriptional studies we were able to confirm the

isolation of FHF and SHF specific CMs from differentiating hESCs. Furthermore, we identify CORIN as a novel cell surface marker of both FHF-like progenitors and CMs. These results provide a platform to investigate *in vitro* cardiovascular development and may facilitate a safe approach for cell therapy in heart disease.

Introduction

In vertebrates, the heart is the first organ to form and plays an essential role in the delivery of oxygen and nutrients to the body throughout development and life. Mammalian heart development is a precise, complex and highly orchestrated process that involves a series of morphogenic and molecular changes that result in generation of multiple cell lineages which ultimately makeup the functioning heart. The heart is formed from the anterior portion of the primitive streak by a small population of mesodermal cells, from which cardiac progenitors arise [1, 2]. The cardiac progenitors then migrate to the lateral plate mesoderm where cell-to-cell interactions are necessary for progenitor specification to a cardiac cell fate [3, 4].

Cardiac specification is also highly influenced by signaling molecules released from adjacent mesoderm and overlying endoderm [5-7]. Inhibitory signals secreted from the axial mesoderm and neural plate delineates the lateral and medial borders of the cardiac region commonly referred to as the cardiac mesoderm [3, 8]. The cardiac mesoderm will then migrate to the midline forming a transient embryonic structure known as the cardiac crescent. During the third week of gestation, cells of the cardiac crescent undergo differentiation and merge along the ventral midline to form a central pump (the primitive heart tube) (Figure 1)[9]. Later in development, asymmetrical looping of the cardiac tube ensures the correct alignment of the atrial and ventricular chambers and subsequent morphogenesis results in the formation of the valves, septa, and conduction system that comprise the mature, functioning heart (Figure 1)[10-12].

The current model of early vertebrate heart development describes two distinct cell sources that contribute to the adult heart [13]. Cells from the first cell source (known as the

first heart field) give rise to the cardiac crescent and myocardial cells of the primitive heart tube (Figure 1). Descendants of the first heart field are found in all chambers of the adult heart. Cells of the second heart field contribute to the primitive heart tube at the arterial pole, which gives rise to the proximal aorta walls and pulmonary trunk in the postnatal heart [14], and the venous pole, which gives rise to the left and right atria and outflow tract (Figure 1) [15].

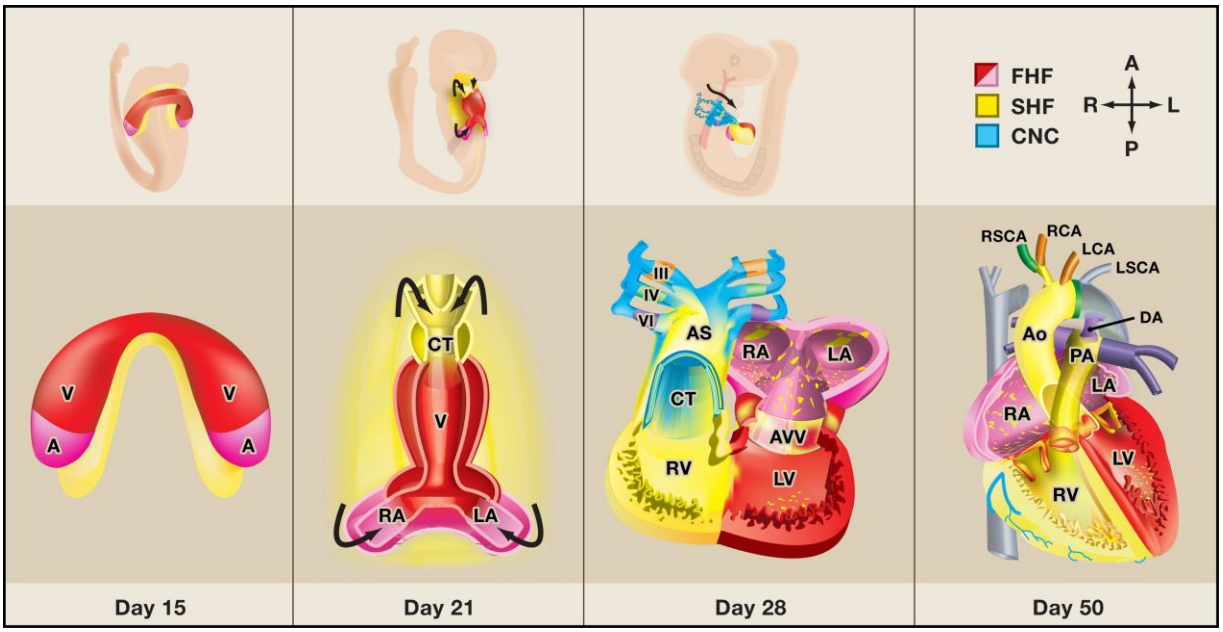


Figure 1. An overview of mammalian heart development. Oblique views of whole embryos and frontal views of cardiac precursors during human cardiac development are shown. (First panel) First heart field (FHF) cells form a crescent shape in the anterior embryo with second heart field (SHF) cells medial and anterior to the FHF. (Second panel) SHF cells lie dorsal to the straight heart tube and begin to migrate (arrows) into the anterior and posterior ends of the tube to form the right ventricle (RV), conotruncus (CT), and part of the atria (A). (Third panel) Following rightward looping of the heart tube, cardiac neural crest (CNC) cells also migrate (arrow) into the outflow tract from the neural folds to septate the outflow tract and pattern the bilaterally symmetric aortic archarteries (III, IV, and VI). (Fourth panel) Septation of the ventricles, atria, and atrioventricular valves (AVV) results in the four-chambered heart. V, ventricle; LV, left ventricle; LA, left atrium; RA, right atrium; AS, aortic sac; Ao, aorta; PA, pulmonary artery; RSCA, right subclavian artery; LSCA, left subclavian artery; RCA, right carotid artery; LCA, left carotid artery; DA, ductus arteriosus. From Srivastava [9].

Using Human embryonic stem cells to study development and a source for cell-based therapies

The *in vitro* differentiation of Human embryonic stem cell (hESCs) recapitulates many aspects of early human development and provides insight to an otherwise inaccessible period of embryogenesis. hESCs have been used to derive cardiac differentiation schemes based on signaling molecules that direct *in vivo* cardiogenesis [16-18]. These studies have been assisted by the generation of several cardiac reporter lines, which have been used to enrich for CPCs as well as their definitive progeny [19-21]. While some investigators have developed strategies to isolate mature CMs using genetic markers such as troponin and myosin heavy chain, others have used early cardiac-related transcription factors such as MESP1 and NKX2-5, both of which are essential for cardiac cell fate specification [22-29].

Despite these advances, current differentiation protocols and cardiac-specific reporter lines yield heterogeneous populations and lack the capacity for the identification, and enrichment, of heart field-specific cardiac progenitors and their progeny. Isolation of heart field-specific CMs is critical for several reasons. First, following a typical myocardial infarction there is extensive and irreversible loss of CMs, primarily localized to the left ventricle (LV) of the heart. The CMs of the LV are predominantly derived from FHF CPCs whereas CMs within the right ventricle (RV) originate from the SHF during early cardiogenesis [30-32]. Although the adult heart consists of CMs from distinct embryonic origins that later localize to specific anatomical regions, it is not known whether their ontogeny influences their physiological function or response to pathological stressors. It is possible that transplantation of heart field-specific CMs may enhance their retention

and integration after transplantation as well as improve cardiac function. Second, there are additional risks associated with the use of a heterogeneous mixture of cells from differentiating hESCs for transplantation. These cells include ventricular and atrial CMs from both heart fields as well as pacemaker cells [33-36]. The contamination of nodal cells could potentially generate a nidus for arrhythmias post-transplantation, which is a major concern that needs to be addressed prior to reaching the clinic. Finally, efficient and reproducible differentiation strategies to generate FHF and SHF CMs will aid modeling and developing treatments for diseases that specifically affect either the left or right regions of the heart.

Attempts to derive heart field-specific cardiac progenitors have been made using ISL1 as a marker for SHF cells [19, 37]. ISL1, while initially expressed throughout the heart, becomes restricted to the SHF during the elongation of the pre-heart tube before being downregulated [19, 37-40]. ISL1 marks a stage-specific, self-renewing cardiac progenitor population capable of giving rise endothelial cells, smooth muscle, and CMs [19]. However, ISL1 also marks cells from other lineages, including β-Islet cells and motor-neurons [41, 42]. As such, cardiac directed hESC-derived $ISL1^{+}$ populations commonly consist of a mixture of CPCs and CM subtypes [43].

In light of this, we took a two-fold approach to identify and purify heart field-specific CMs for cell transplantation. First, we generated a FHF-specific, TBX5-TdTomato hESC reporter line. *In vivo, TBX5* is predominately expressed in the primitive posterior heart tube, marking progenitors of the left ventricle and atria [44-50]. TBX5 has also been shown to interact with NKX2-5, an important transcription factor in cardiac development [51]. We further developed this system into a double hESC reporter line (*TBX5*$^{TdTomato/W}$*/NKX2-*

5$^{eGFP/W}$) that offers a more precise and versatile platform for isolating heart field-specific CMs. Utilizing these cell lines, we were able to isolate and characterize CMs derived with molecular identities that resemble primary and secondary heart fields using RT-qPCR, immunocytochemistry, RNA-sequencing, *in vivo* cell transplantation, and electrophysiological studies. These reporter systems were used to optimize the *in vitro* differentiation of CMs that display similar characteristics to FHF and SHF CMs as well as nodal cells. Finally, we discovered that CORIN is a novel surface marker that can be used to selectively isolate FHF-like CMs, alleviating the requirement for genetically modified hESC lines. Overall, our studies demonstrate that heart field-specific CMs can be prospectively isolated from differentiating hESCs. This is an essential advancement for *in vitro* developmental studies as well as cell transplantation therapies for heart disease.

Methods

Maintenance of hESC

Feeder-free culture of hESC lines were carried out on Matrigel matrix (Corning) or Geltrex (Thermo Fisher Scientific) coated dishes with mTeSR1 medium (STEMCELL Technologies) and was changed every day. Cells were passaged at 80-90% confluence using ReLeSR (STEMCELL Technologies).

Generation of reporter cell lines

To generate a non-perturbing knock-in reporter cell line for TBX5, the T2A self-cleaving peptide followed by TdTomato gene were incorporated in-frame into the last exon (Exon 9) of *TBX5* without affecting its 3'UTR using TALENs. The *TBX5*-T2A-TdTomato donor vector was built to contain left- and right homology arms (800bp-1000bp) that flank the genomic cleavage site in the *TBX5* locus. Southern blotting was performed using probes and enzymes described in Table S1. Karyotyping was performed by WiCell Research Institute. The H9-*TBX5*$^{\text{Tdtomato}}$*TNNT2*$^{\text{copGFP}}$ and *TBX5*$^{\text{Tdtomato}}$*MYL2*$^{\text{copGFP}}$ reporters were constructed by transducing the H9-*TBX5*$^{\text{TdTomato}}$ cell line with *TNNT2*-copGFP and *MYL2*-copGFP lentiviruses according to the manufacturer's instructions. The pre-packaged lentiviral preparations of pGreenZeo-*TNNT2*-copGFP and pGreenZeo-*MYL2*-copGFP reporters were obtained from System Biosciences (SBI). The HES3-*NKX2-5*$^{eGFP/W}$ reporter cell line was generously provided by E. Stanley and A. Elefanty (Monash University, Victoria, AU) and was generated by targeting the eGFP coding sequence to the *NKX2-5* locus of HES3 cells using previously described protocol [21]. The HES3-*TBX5*$^{\text{TdTomato/W}}$/*NKX2-5*$^{\text{eGFP/W}}$ double reporter line was generated by using

the same *TBX5* targeting strategy to insert *TBX5*-T2A-TdTomato into the HES3-*NKX2-5*$^{eGFP/W}$ reporter line.

Differentiation of hESC

Monolayer differentiations of cardiomyocytes were carried out using the GiWi strategy as previously described [52]. Briefly, H9-*TBX5*^{TdTomato} hESCs were grown on Matrigel or Geltrex to 90% confluency. The cells were treated with 12µM CHIR99021 (Tocris Biosciences) in RPMI1640 media (Thermo Fisher) supplemented with B27 minus insulin (RPMI B27-ins; Thermo Fisher Scientific) on Day 0. Exactly 24 hours later (Day 1), the media was replaced with RPMI B27-ins. On Day 3 and Day 5, the media was replaced with RPMI B27-ins with 5µM IWP2 (Tocris Biosciences) and RPMI B27-ins respectively. From Day 7 onwards cells were maintained in RPMI media supplemented with B27 (RPMI B27; Thermo Fisher Scientific) until analysis.

HES3-*TBX5*$^{TdTomato/W}$/*NKX2-5*$^{eGFP/W}$ hESCs were grown on Geltrex to 90% confluency then harvested as single cell suspension using Accutase (Thermo Fisher Scientific) and resuspended in mTeSR1 containing 10µM ROCK inhibitor Y-27632 (Tocris Biosciences). Cells were counted using a Countess II Automated Cell Counter (Thermo Fisher) and re-plated onto Geltrex coated plates at 1.3 x 10^5 cells/cm^2 for FHF (TBX5$^+$NKX2-5$^+$), 2.0 x 10^5 cells/cm^2 for SHF (TBX5$^-$NKX2-5$^+$) and 3.0 x 10^5 cells/cm^2 for nodal (TBX5$^+$NKX2-5$^-$ SIRPα$^+$CD90$^-$) in mTeSR1 containing 10µM ROCK inhibitor Y-27632 (day -2 of differentiation). At day -1 media was changed to mTeSR1. At day 0 media was changed to RPMI B27-ins containing CHIR99021 (10µM for FHF, 6µM for SHF and 12µM for nodal). After 24 hr (day 1) media was changed to RPMI B27-ins until

day 3. On day 3 of differentiation cells were changed to RPMI B27-ins containing 5μM IWP2. At day 5, media was changed to RPMI B27-ins until day 7 when media was switched to RPMI B27. Cells were maintained in this media and changed every 3 days thereafter.

Flow cytometry and cell sorting

Differentiated hESCs were dissociated with TrypLE Express (Thermo Fisher) for 3-4 min at 37°C to form a single cell suspension. Antibody labelling of cells was performed using anti-VCAM1 APC (Biolegend), anti-SIRPα PE-Cy7 (Biolegend), anti-Tra1-81 Alexa Fluor 594 (Biolegend), anti-CD90 APC (Biolegend), anti-Podoplanin PE-cy7 (Biolegend), anti-CORIN (KAN Research Institute Inc.), anti-TNNT2 BV421 (BD Bioscience) and anti-mouse IgG APC (Biolegend). Anti-mouse IgG APC was used at a dilution of 1:50, all other antibodies were used at a dilution of 1:100. Detailed antibody information is provided in Table S3. All antibodies were diluted and incubated in FACS buffer (2% FBS, 1% BSA, 2mM EDTA) containing 10μM ROCK inhibitor Y-27632. Samples were incubated with antibodies for 30-60 min on ice and washed using FACS buffer. Control stains using non-specific antibodies (IgG) were performed. Cells were sorted using a FACS-ARIA (BD Biosciences) into RPMI B27 with 10μM ROCK inhibitor Y-27632. Flow analysis was performed using a Fortessa flow cytometer (BD Biosciences) and analyzed using FlowJo software (Tree Star Inc).

For intracellular staining of TNNT2, cells were first stained for CORIN as described above, washed and incubated with anti-mouse IgG APC secondary for 30 min on ice in

FACS Buffer, then stained for TNNT2 using the BD Cytofix/Cytoperm fixation and permeabilization kit (BD Biosciences) according to the manufacture's instructions.

RNA-sequencing

Total RNA of the cells was isolated using TRIZOL LS Reagent (Thermo Fisher), chloroform extraction and isopropanol precipitation followed by further purification using RNeasy Micro kit (Qiagen). The quality of the RNA was assessed by Agilent 2200 Tapestation. For library preparation, total RNA was fragmented and subjected to cDNA conversion, adapter ligation and amplification using KAPA Stranded RNA-Seq Library Preparation Kit according to the manufacturer's instructions. The final library was quantified using Agilent 2100 Bioanalyzer to evaluate its integrity. The deep sequencing for 2x150 bp paired-end reads was performed using Illumina HiSeq 2500 Sequencing System.

For sample analysis, RNA-seq data was mapped to the reference genome (GRCh38) with OLego version 1.1.5 [53] and normalized by using TPM (Transcripts per millions) analysis. Total number of reads mapped to a known transcript annotation was estimated using featureCounts version v1.5.0-p2 [54]. Expression levels for each transcript were determined by normalizing the counts returned by featureCounts using custom Perl scripts. Normalized expression levels for each transcript were determined by transforming the raw expression counts to TPM following log2 scaling. Gene Ontology (GO) enrichments were computed using DAVID Bioinformatics Resources v6.8. RStudio was used to run custom R scripts to generate heatmaps using 'heatmaply' package.

Quantitative real-time RT-PCR

Total RNA was extracted from cells using RNeasy mini or micro kit (Qiagen) according to the manufacturer's instructions. Purified RNA was subjected to reverse transcription as described in the iScript cDNA Synthesis Kit (Bio-Rad Laboratories). Following cDNA synthesis, expression level of the genes was quantified by quantitative Real-Time PCR (qPCR) using iTaq universal SYBR Green supermix on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratories). The quantification was analysed using the C_T (threshold cycle) values. The expression of the target genes was normalized to *GAPDH* and the relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method. Primers used in this study are in Table S2.

Immunofluorescent analyses

Cells cultured on Geltrex coated glass coverslips were fixed in 4% PFA in PBS for 10 minutes at RT followed by PBS washings. Explanted hearts were fixed in 4% PFA overnight at 4°C, then cryoprotected in 30% sucrose solution overnight at 4°C and embedded in Tissue-Plus O.C.T. compound (Fisher Scientific). The tissue blocks were frozen at -80°C, after which the frozen blocks were cryo-sectioned using Leica CM 1860 cryostat (Leica Biosystems). For each block, 8µm thick cross sections were collected on Superfrost Plus Microscope slides (Fisher Scientific), which were later used for immunohistochemistry staining procedures.

For staining, fixed cells or tissue sections were permeabilized with 0.2% Triton X-100 in PBS for 10 minutes prior to blocking non-specific binding with 10% normal serum (from the same host species as the labeled secondary antibody) in PBST (PBS with 0.1% Tween-20) for 30 minutes. Cells and sections were incubated with primary antibodies

diluted in blocking buffer at 4°C overnight, washed 3 times in PBST, then stained with secondary antibodies diluted in blocking buffer at RT for 1 hour followed by 3 washes in PBST. The cell nuclei were counterstained with DAPI using VECTASHIELD Antifade Mounting Medium with DAPI (Vector Laboratories). The stained cells and tissue sections were imaged with a Leica TCS SP5 microscope using LAS X software (Leica Biosystems) or a LSM 880 with Airyscan Confocal Microscope using ZEN software (Carl Zeiss Microscopy).

The following primary antibodies were used: Rabbit anti-Cardiac Troponin T (Abcam; 1:400), mouse anti-TBX5 (Santa Cruz; 1:100), goat anti-ISL1 (R&D Systems; 1:20), mouse anti-SHOX2 (Abcam; 1:200), mouse anti-cardiac Troponin T (Thermo Fisher; 1:400), rabbit anti-Oct4 (EMD Millipore; 1:200), mouse anti-Nanog (EMD Millipore; 1:500), mouse anti-mitochondria (Abcam; 1:800), rabbit anti-CD31 (Abcam; 1:100), rabbit anti-CNN1 (Sigma; 1:200), rabbit anti-Cx43 (Cell Signaling Technology; 1:200) and rabbit anti-CORIN (Abcam; 1:100). Detailed antibody information is provided in Table S3.

The following secondary antibodies were used: Goat anti-rabbit Alexa Fluor 488 (Thermo Fisher; 1:150), goat anti-mouse Alexa Fluor 488 (Thermo Fisher; 1:150), goat anti-rabbit Alexa Fluor 594 (Thermo Fisher; 1:150), goat anti-mouse Alexa Fluor 594 (Thermo Fisher; 1:150) rabbit anti-mouse Alexa Flour 647 (Thermo Fisher; 1:150), donkey anti-goat Alexa Flour 647 (Thermo Fisher; 1:150).

Ischemia/Reperfusion injury and cell transplantation in mice

Experiments were performed in 8-10-week-old male non-obese diabetic/SCID mice (NSG) with common γ-chain KO. Mice were anesthetized by intraperitoneal injection

of ketamine/xylazine. Mice were intubated for positive pressure ventilation with oxygen-enriched room air during the surgical procedure. After left thoracotomy between ribs four and five, the pericardium was opened, and 8-0 suture were placed around the left anterior descending artery (LAD) which is 1-2 mm from the tip of the normally positioned left atrium and the suture was tightened to occlude the vessel for 45 min before being released. The chest was closed in layers using 5-0 Vicryl suture and ventilation was maintained until sufficient spontaneous breathing occurs. Seven days after injury, cells were transplanted at three sites into the border zone under direct visualization. Approximately $1.2x10^6$ cells were resuspended in growth factor reduced Matrigel (Corning) and RPMI B27 (1:1) then injected at each site using 27-guage needle. $TBX5^+$ and $TBX5^-$ cells were sorted on day 7 of differentiation and re-cultured in RPMI B27 media for two additional days before transplantation. $TBX5^+NKX2\text{-}5^+$, $TBX5^-NKX2\text{-}5^+$, and $CORIN^+$ cells were sorted on day 10 of differentiation and re-cultured in RPMI B27 media for three additional days before transplantation. Mice were sacrificed, and hearts were harvested for sectioning and immunohistochemistry either 7-8 weeks post-transplantation ($TBX5^+$ and $TBX5^-$) or 5-6 weeks post-transplantation ($TBX5^+NKX2\text{-}5^+$, $TBX5^-NKX2\text{-}5^+$, and $CORIN^+$).

Teratoma formation assay

Adult male NSG mice (8-10 weeks old) were anesthetized with isoflurane and after a transverse incision to expose the kidney, $5x10^5$ cells resuspended in growth factor reduced Matrigel and RPMI B27 (1:1) were injected directly underneath the capsular membrane of the kidney using a 26-gauge Hamilton syringe. The kidney was then gently repositioned into the abdominal cavity and the muscle and skin was sutured. Four weeks after transplantation the kidney was removed and processed for histology analysis.

Patch clamp

For electrophysiological characterization using patch clamp cultured cells were dissociated and respective cell populations were isolated by FACS as described above. Isolated cells were resuspended in RPMI B27 supplemented with 10μM ROCK inhibitor Y-27632 at 1-2x10^6 cells/ml. Drops of 25μl of this cell suspension were applied to glass cover slips (5mm) that were pre-coated with Geltrex in 6-well plates. The cells were incubated in the 25μl volume for 8-12 hr to facilitate cell attachment. The dishes were then flooded with 2ml of RPMI B27 supplemented with 10μM ROCK inhibitor Y-27632. ROCK inhibitor was removed after 24 hrs and the media was changed every 2 days. Cultures were used for patch clamp recordings 4–14 days following plating. Action potentials were measured using standard patch- clamp techniques in current-clamp modes, (Axopatch 200B, Axon Instruments). Voltages were recorded with 2 KHz sampling rate (DigiData 1200, Axon Instruments) and analyzed with Clampfit software (Axon Instruments). Borosilicate glass microelectrodes were used with tip resistances of 2–5 MΩ when filled with pipette solution.

Monolayer Optical Mapping

For high-resolution optical mapping of action potential (AP) propagation, cultured cells were dissociated and respective cell populations were isolated by FACS as described above. Isolated cells were suspended in RPMI B27 supplemented with ROCK inhibitor Y-27632 (10μM) at 20-22x10^6 cells/ml. Drops of 25μl of this cell suspension were applied to Geltrex-coated 5mm coverslips (5x10^5 cells/coverslip). The cells were incubated in the 25μl volume for 8-12 hr to facilitate cell attachment. The dishes were then flooded with 2ml of RPMI B27 supplemented with 10μM ROCK inhibitor Y-27632.

ROCK inhibitor was removed after 24 hrs and the media was changed every 2 days. On day of optical mapping recordings, once contractions were observed, cells were stained with voltage-sensitive dye, Di-8-ANEPPS (40μM; Thermo Fisher) and washed with normal Tyrode solution three times. Optical AP recording were made using MiCAM-Ultima CMOS camera (100 × 100 pixels, 16μm/pixel) at 500 frames per second (fps). First, spontaneously occurring APs were recorded, followed by overdrive electrical pacing (1Hz, 0.5Hz). The basic cycle length was gradually decreased until loss of 1:1 capture or re-entry was induced.

Chronotropic response to β-adrenergic agonist and antagonist

Day 10 FACS isolated cells were re-plated on Geltrex coated dishes for 7-10 days. Basal beating rates were counted using a light microscope then 1μM Isoproterenol (Sigma) was added and incubated at 37°C for 5 min before beating rates were counted again. 10 μM Propranolol (Sigma) was then added, incubated at 37°C for 5 min and beating rate counted again.

Animal Use and Care

Animal housing, handling, and procedures were approved by and carried out in accordance with guidelines set by the UCLA Animal Research Committee (ARC), UCLA Institutional Animal Care and Use Committee (IACUC), and the NIH Guide for the Care and Use of Laboratory Animals.

Human fetal tissue

Anonymized Human fetal heart samples were obtained from the UCLA Translational Pathology Core Laboratory as approved by UCLA IRB.

Statistics

All data are represented as mean ± standard error of mean (SEM). Indicated sample sizes (n) represent biological replicates and individual tissue samples. No statistical method was used to predetermine the samples size. Due to the nature of the experiments, randomization was not performed and the investigators were not blinded. Statistical significance was determined by using Student's t test (unpaired, two-tailed) in GraphPad Prism 6 software. Results were considered to be significant at $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), and $p < 0.0001$ (****). All statistical parameters are reported in the respective figures and figure legends.

Results

Generation and validation of a FHF-specific cardiac progenitor hESC reporter line

TBX5 is predominately expressed in the primitive posterior heart tube, marking progenitors of the LV and atria, corresponding to the FHF [44-50]. To enrich for FHF-progenitors from differentiating hESCs, we introduced the TdTomato (TdT) gene into the *TBX5* locus of H9 hESCs (*TBX5*$^{TdTomato/W}$) by TALEN-mediated homologous recombination. To avoid haploinsufficiency, the TALEN targeting strategy was designed such that the TdTomato gene was fused with exon 9 (last exon) of *TBX5* via a T2A peptide, which allowed accurate reporter expression while maintaining biallelic expression of TBX5 (Figure 2A). The correctly targeted clone was identified by PCR analysis and validated by Southern Blotting using *TBX5*-specific probes (Figure 2B). The cells exhibited a normal karyotype (Figure 2C), successfully differentiated into teratomas contributing to all three germ layers when transplanted under the kidney capsule in NSG mice (Figure 2D) and expressed pluripotency markers after several passages (Figure 2E and 2F).

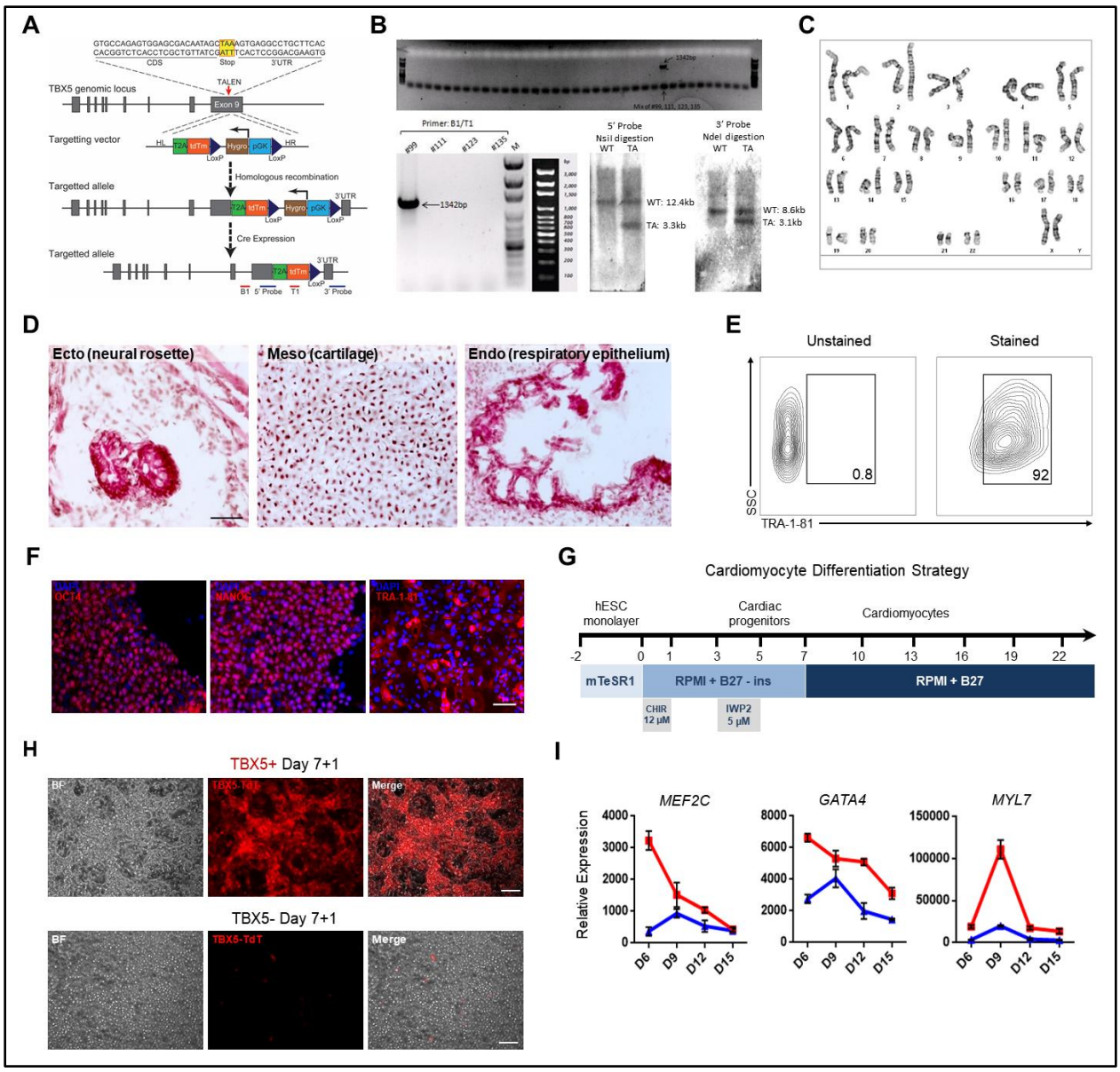


Figure 2. Generation and validation of the H9-*TBX5*$^{TdTomato/W}$ hESC line. (A) Schematic of the TALEN targeting strategy used to insert TdTomato into the *TBX5* locus. (B) PCR screening using B1/T1 primers to identify correctly targeted *TBX5*-TdT clones. Southern blotting of positive H9-*TBX5*^{TdTomato} clone with 5'arm and 3'arm probe of *TBX5* to validate the targeting of the *TBX5* gene. (C) Karotyping analysis of H9-*TBX5*^{TdTomato} cells. (D) Teratoma formation of H9-*TBX5*^{TdTomato} hESCs when transplanted into the kidneys of NSG mice depicting the three germ layers (Scale bar = 75µm). (E and F) Flow cytometry analysis and immunofluorescence staining for pluripotency genes (Scale bar = 50µm). (G) Schematic representation of the cardiomyocyte differentiation strategy used. (H) Brightfield (BF), TdT (red) and overlayed photomicrographs of TBX5+ and TBX5- cells 1-day post sorting (Day 7+1) (Scale bar = 75µm). (I) Relative gene expression of cardiac genes in TdT+ and TdT- populations at days 6-15 during differentiation normalized to undifferentiated hESCs (n=3, Error bars represent SEM). Scale bars = 100µm.

In vitro differentiation reveals enrichment of FHF progenitors in TBX5^{+} cells

We first explored the lineage commitment of TBX5-marked cells from differentiating H9-*TBX5*$^{TdTomato/W}$ hESC line using a monolayer cardiac differentiation protocol (Figure 2G) described previously [52]. Flow cytometry analysis beginning at day 4 of differentiation showed progressive increase in the expression of TdT which was widely expressed by day 7-12 of differentiation (Figure 3A). Further, TdT^{+} cells were readily sorted using fluorescence activated cell sorting (FACS) (Figure 2H and 3A). Gene expression analysis on FACS isolated populations demonstrated that the TdT^{+} population had significantly higher levels of *TBX5* mRNA than TdT^{-} fractions at all time points examined (Figure 3B). Furthermore, TBX5^{+} (TdT^{+}) cells showed enriched expression of mesoderm and cardiac progenitor transcripts, including *NKX2-5*, *MEF2C* and *GATA4* as well as cardiac sarcomeric genes such as *TNNT2*, *ACTN2*, *MYL2* and *MYH7* (Figure 2I and 3B). In contrast, the TBX5^{-} population expressed higher levels of *ISL1* and lower levels of contractility genes during differentiation (Figure 2I and 3B), suggesting that perhaps TBX5^{-} cells generated by this differentiation protocol may contain SHF progenitors in addition to other cells of mesoderm origin. Flow cytometry at day 7 revealed that nearly all TBX5^{+} cells (>97%) expressed the CPC marker, SIRPα, with 23% also co-expressing VCAM1 (vs 90% and 7% respectively in TBX5^{-} cells), providing evidence that TBX5^{+} cells preferentially commit to a CM phenotype (Figure 3C) [21, 55]. Interestingly, when sorted on day 7 and re-cultured under cardiac conditions for an additional 7 days (day 7+7), the majority of TBX5^{+} cells proceed to upregulate VCAM1 and express TNNT2 (Figure 3C and 3D), with an increased propensity to form spontaneously contracting

monolayers. These data demonstrate that H9-*TBX5*$^{TdTomato/W}$ faithfully report expression of the endogenous gene.

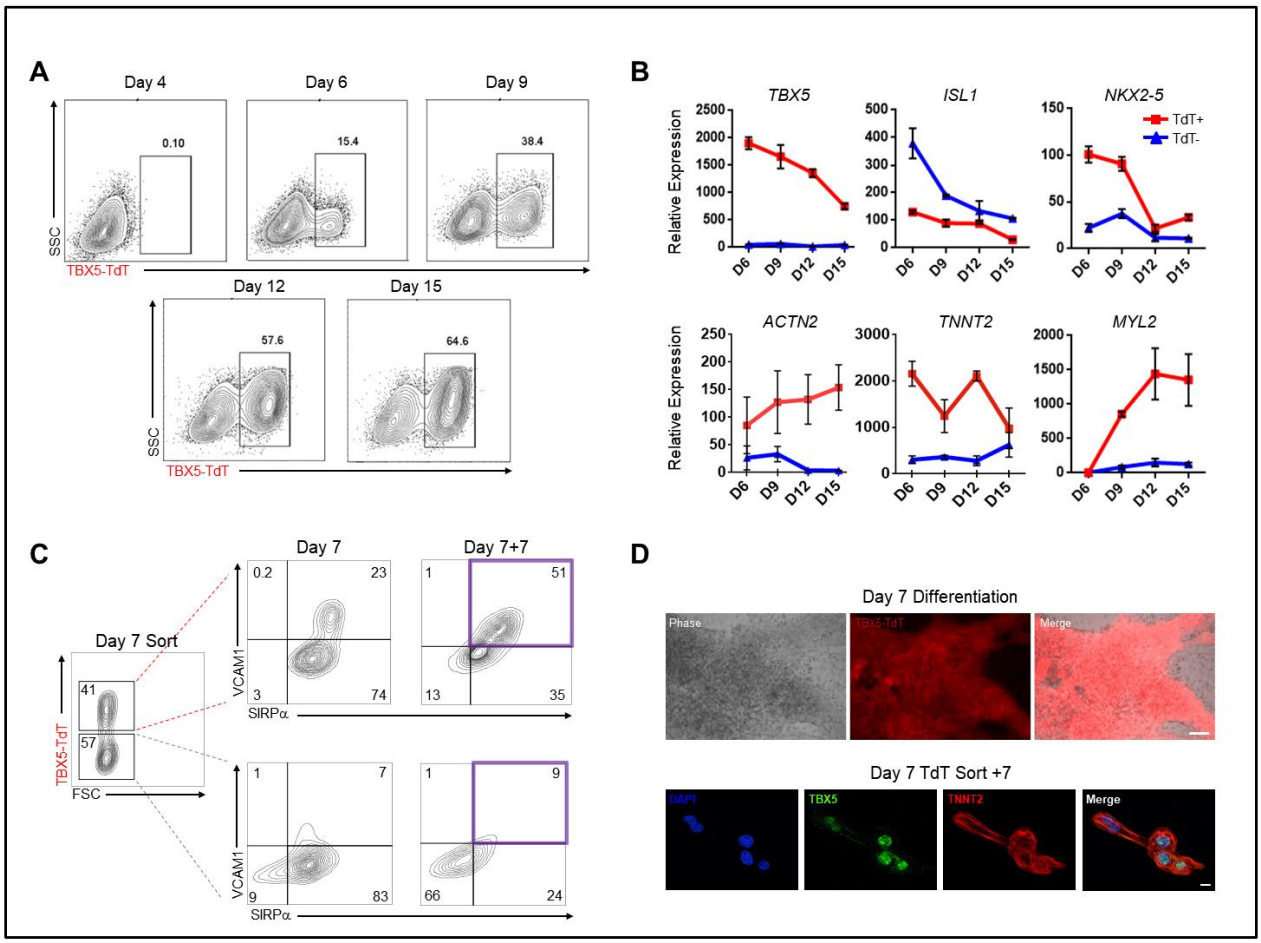


Figure 3. Characterization of the H9-*TBX5*$^{TdTomato/W}$ hESC line. (A) Flow cytometric analysis of TdTomato (TdT) expression in differentiating hESCs. (B) Relative gene expression of cardiac genes in TdT+ and TdT- populations at days 6-15 during differentiation normalized to undifferentiated hESCs (n=3, Error bars represent SEM). (C) FACS plots showing percentage of TdT+ and TdT- cells on day 7 and the percentages of SIRPα+ and VCAM1+ cells within the TdT+ and TdT- populations on day 7 and day 7+7. (D) Phase-contrast (Phase), TdT (red), and overlayed photomicrograph of day 7 differentiated H9-*TBX5*^{TdTomato} cells (top) (Scale bar = 75µm). Immunofluorescence expression of TBX5 (green) and the cardiomyocyte marker (TNNT2; red) 7 days post sorting (Day 7+7) (Bottom) (Scale bar = 10µm).

We then determined whether TBX5^{+} and TBX5^{-} progenitors are capable of giving rise to other cardiac cell types such as smooth muscle (SMC) and endothelial cells (EC). Gene expression analysis on FACS isolated populations at various timepoints of differentiation showed expression of SMC (*CNN1*) and EC (*CD31*) genes in both TBX5^{+} and TBX5^{-} populations (Figure 4A). Immunostaining on TBX5^{+} and TBX5^{-} sorted cells confirmed both TBX5^{+} and TBX5^{-} progenitors can give rise to clusters of SMC and EC (Figure 4B). Together, these data suggest that while expression of TBX5 during early differentiation of hESCs may favor generation of FHF CMs, both TBX5^{+} and TBX5^{-} progenitors remain multi-potent with the capacity to differentiate into SMCs and ECs *in vitro*.

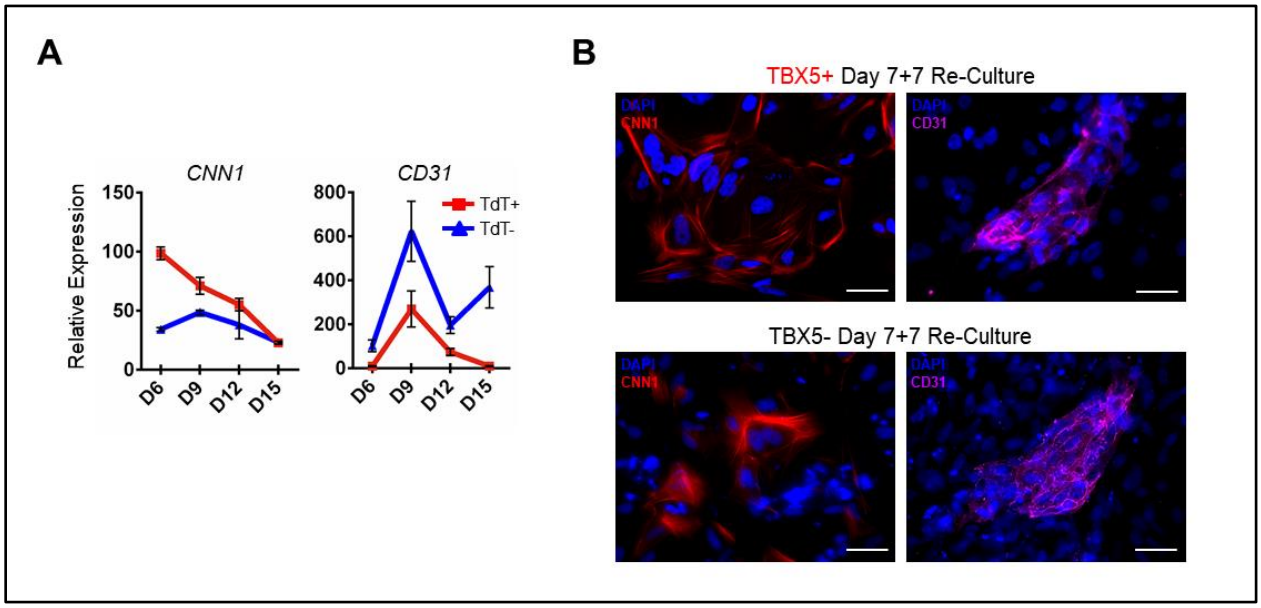


Figure 4: Endothelial and Smooth muscle contribution of TBX5 cells. (A) Gene expression analysis of TBX5^{+} and TBX5^{-} populations at various timepoints of differentiation for SMC (*CNN1*) and EC (*CD31*) genes (n=3, Error bars represent SEM). (B) Immunofluorescence analysis of SMC and EC markers (CNN1 and CD31, respectively) in TBX5+ and TBX5- cells sorted at day 7 and re-cultured in cardiac culture conditions for 7 days (Day 7+7) (Scale bars = 50µm).

TBX5$^+$ cells show engraftment and differentiation into cardiomyocytes when transplanted into the injured left ventricle of mice

To evaluate the *in vivo* engraftment and differentiation potential of TBX5$^+$ and TBX5$^-$ progenitors, day 7 sorted fractions were cultured for an additional 2 days then transplanted into the left ventricular peri-infarct area of NSG mice 7 days after ischemia-reperfusion (I/R) injury. Human specific mitochondrial staining revealed many clusters of TBX5$^+$ cells (Figure 5A) whereas fewer TBX5$^-$ cells were retained in the transplanted hearts (Figure 5B). Much of the transplanted TBX5$^+$ and TBX5$^-$ grafts showed a TNNT2$^+$ CM phenotype with expression of the gap junction protein, Cx43, within the grafts, but not between the host CMs and transplanted cells (Figure 5C and 5D). Examination of the TNNT2 staining revealed that the grafted CMs displayed disorganized sarcomeric structure compared to the host CMs (Figure 5E and 5F). In contrast to our *in vitro* data, we did not observe CD31 or CNN1 within the grafts (Figure 5C and 5D), suggesting that the TBX5$^+$ and TBX5$^-$ fractions may have a preferential differentiation towards CMs *in vivo*. Additionally, we did not observe any teratoma formation in either of the transplanted populations. Collectively, these results indicate that hESC-derived TBX5$^+$ progenitors exhibited greater survival after transplantation into the injured LV of mice, where they predominantly give rise to CMs without any evidence of structural integration into the host myocardium.

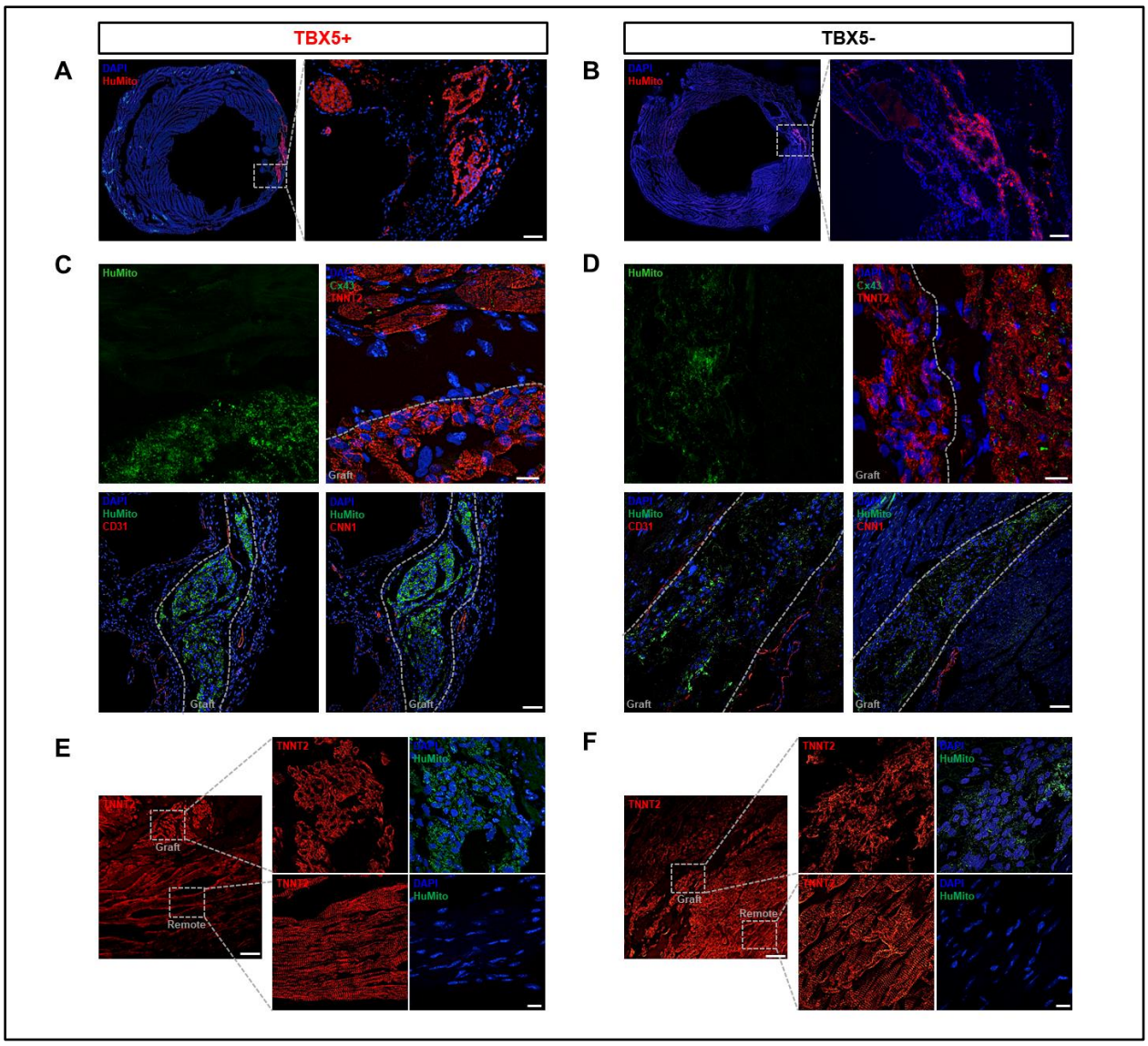


Figure 5. *In vivo* engraftment and differentiation potential of TBX5+ and TBX5- progenitors. Immunohistochemical analysis of engrafted (A) TBX5+ or (B) TBX5- cells stained with a human specific mitochondria antibody. All cells were visualized using DAPI. Dashed box shows magnified area on right (Scale bar = 100µm). (C and D) Analysis of graft areas for the presence of gap junction protein Cx43, endothelial cells (CD31) and smooth muscle cells (CNN1) in TBX5+ and TBX5- transplanted hearts. Cardiomyocytes are marked with TNNT2 and the presence of human cells is shown by human mitochondria staining. Graft areas are marked with dashed lines (Scale bars top = 20µm, bottom = 100µm). (E and F) Analysis of TNNT2 sarcomeric structure of transplanted TBX5+ and TBX5- cells (Scale bar = 100µm). Dashed boxes outline magnified graft and host areas shown on the right (Scale bars = 20µm).

Transcriptional profiling of TBX5 cells at different developmental stages

Our results showed the ability of both TBX5^{+} and TBX5^{-} progenitors to give rise to CMs. To identify potential differences between these two populations at the transcriptomic level, we performed detailed RNA sequencing analyses at early and late stages of cardiac differentiation. We used TNNT2^{copGFP} and MYL2^{copGFP} lentiviral reporter constructs to generate dual H9-*TBX5*^{Tdtomato}/*TNNT2*^{copGFP} and -*TBX5*^{Tdtomato}/*MYL2*^{copGFP} reporter lines, which enabled the isolation of early stage TNNT2^{+} CMs at Day 14 and later stage MYL2^{+} ventricular CMs at Day 30 of differentiation (Figure 7A). These cell lines, along with the previously generated H9-*TBX5*^{TdTomato} single-reporter, were differentiated and FACS isolated on days 7, 14, and 30 and compared to undifferentiated hESCs (day 0) as a baseline (Figure 6A and 7B).

We paid particular attention to the gene expression patterns of CMs (TNNT2^{+} and MYL2^{+}) derived from both TBX5^{+} and TBX5^{-} (Figure 6B). Principal component analysis (PCA) confirmed the consistency in overall gene expression changes across samples within each group (Figure 6C). While a clear distinction in the transcriptomes between day 0 hESCs and the differentiated populations was observed, overall differences between early and late stage CMs appeared to be independent of TBX5 expression (Figure 6C, 7C and 7E). However, in gene groups related to cardiac development and maturation there were key differences between TBX5^{+} and TBX5^{-} populations and their derivatives (Figure 6D). Both TBX5^{+} and TBX5^{-} CM populations showed increased expression for cardiac contraction, ventricular-, and atrial-related genes. Notably, expression of SHF markers such as *ISL1* and *TBX1* were enriched in TBX5^{-} populations, suggesting the possibility that TBX5^{-} CMs may consist of SHF-like CMs (Figure 6D and

7G). These data suggest that the TBX5 reporter line may be used as a tool for isolating heart field-specific populations of hESC-derived CMs. Gene ontology (GO) analysis of the top differentially expressed genes between day 7 $TBX5^+$ and $TBX5^-$ populations revealed that $TBX5^+$ cells are more enriched for cardiac development and contraction pathways (Figure 6E). Additional differences were observed at later stages of development (Figure 7D and 7F). $TBX5^+TNNT2^+$ CMs at day 14 had a gene expression profile that included pathways related to energy production and ATP synthesis, suggestive of a possible shift toward gene programs required for initiation of contraction by these immature CMs (Figure 7D). By day 30, more mature $TBX5^+MYL2^+$ ventricular CMs showed an even greater enrichment of contraction and cardiac development genes (Figure 7F).

Due to the lack of robust surface markers for hESC-derived CMs, we sought to identify novel surface markers within our RNA-sequencing dataset that are distinctly expressed on CPCs and continue their expression when differentiated to CMs. Using the DAVID Bioinformatics Resource, we identified the top 10 highly expressed surface markers from the $TBX5^+$, $TBX5^+TNNT2^+$, and $TBX5^+MYL2^+$ populations, 7 of which were common to these groups. Expression of these particular surface markers were low in the hESC, $TBX5^-$, $TBX5^+TNNT2^-$, and $TBX5^+MYL2^-$ populations, indicating the selectivity of these markers for hESC-derived CMs (Figure 6F). We discovered 3 potential surface markers that consistently ranked high (top 3) when we compared their expression ratios at the 3 differentiation time points: *VCAM1*, *CRYAB*, and *CORIN* (Figure S3H). *CORIN* ranked in the top 3 expression ratios at all 3-time points and appeared to be the most suitable surface marker for differentiating CMs since its expression levels remained high throughout CM maturation.

Collectively, gene expression analysis of hESC-derived cardiac cells at different stages of development suggests that while TBX5 expression leads to an enrichment of FHF-like CMs, $TBX5^-$ populations also generate cardiac cells that resemble SHF CMs at the transcriptional level. Furthermore, Corin was identified as a top surface marker candidate for prospective identification and isolation of hESC-derived FHF-like CMs.

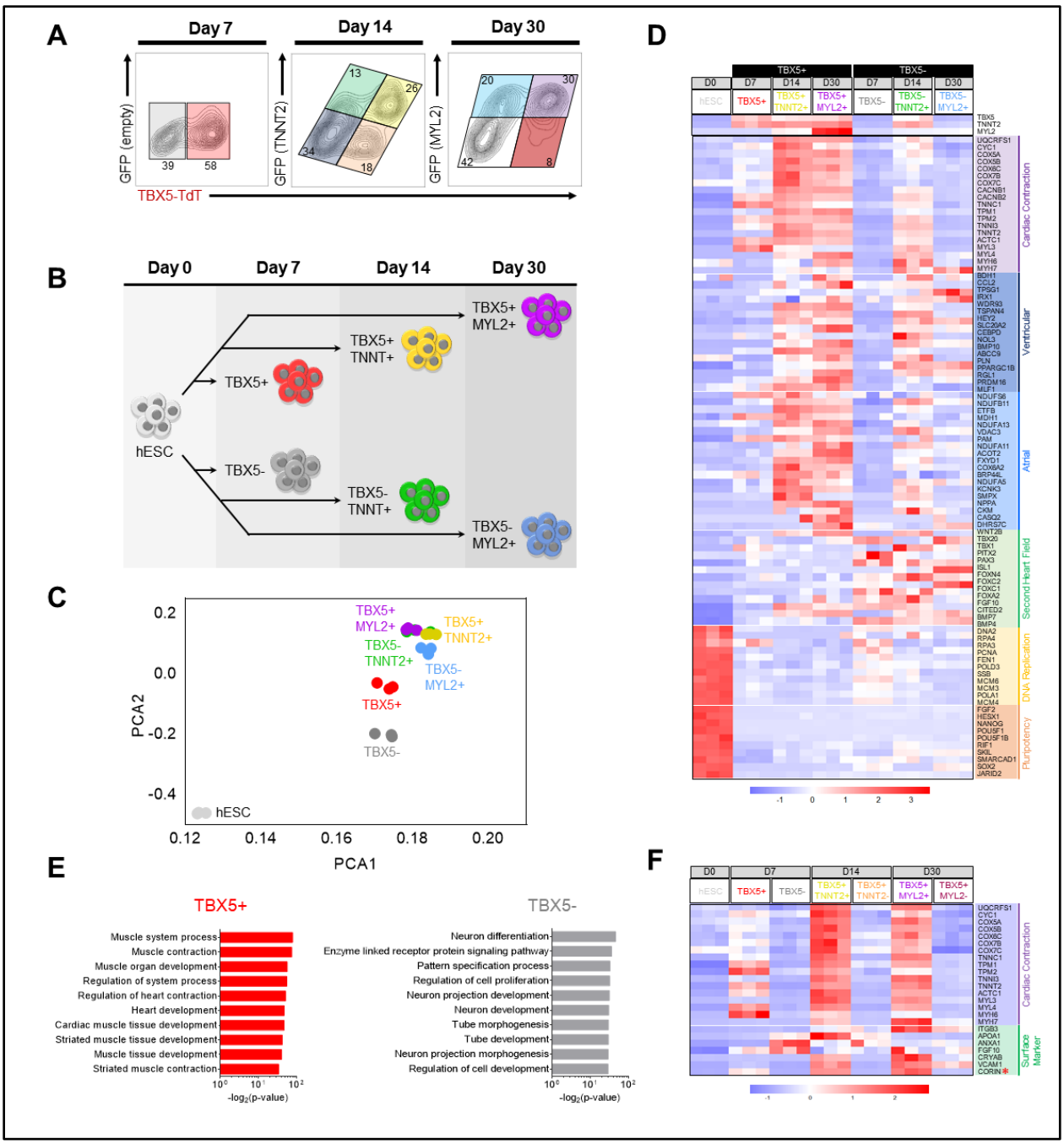


Figure 6. Transcriptional profiling of TBX5+ and TBX5- populations at different developmental stages. (A) Representative flow cytometric plots showing the cell sorting strategy used for isolating distinct populations at day 7, 14 and 30 of differentiation. (B) Schematic showing TBX5+ and TBX5- derived cardiomyocyte populations at varying timepoints of differentiation. (C) PCA analysis of TBX5+ and TBX5- derived cardiomyocyte populations. (D) Heatmap comparing cardiac-related and pluripotency gene expression of TBX5+ and TBX5- derived cardiomyocyte populations. (E) GO analysis of the top differentially expressed genes between TBX5^{+} and TBX5^{-} populations at day 7 of cardiac differentiation. (F) Heatmap showing co-expression of cardiac contraction and surface marker genes in TBX5+ and TBX5- cells at day 7 and TBX5+ cardiac and non-cardiac populations at day 14 and 30.

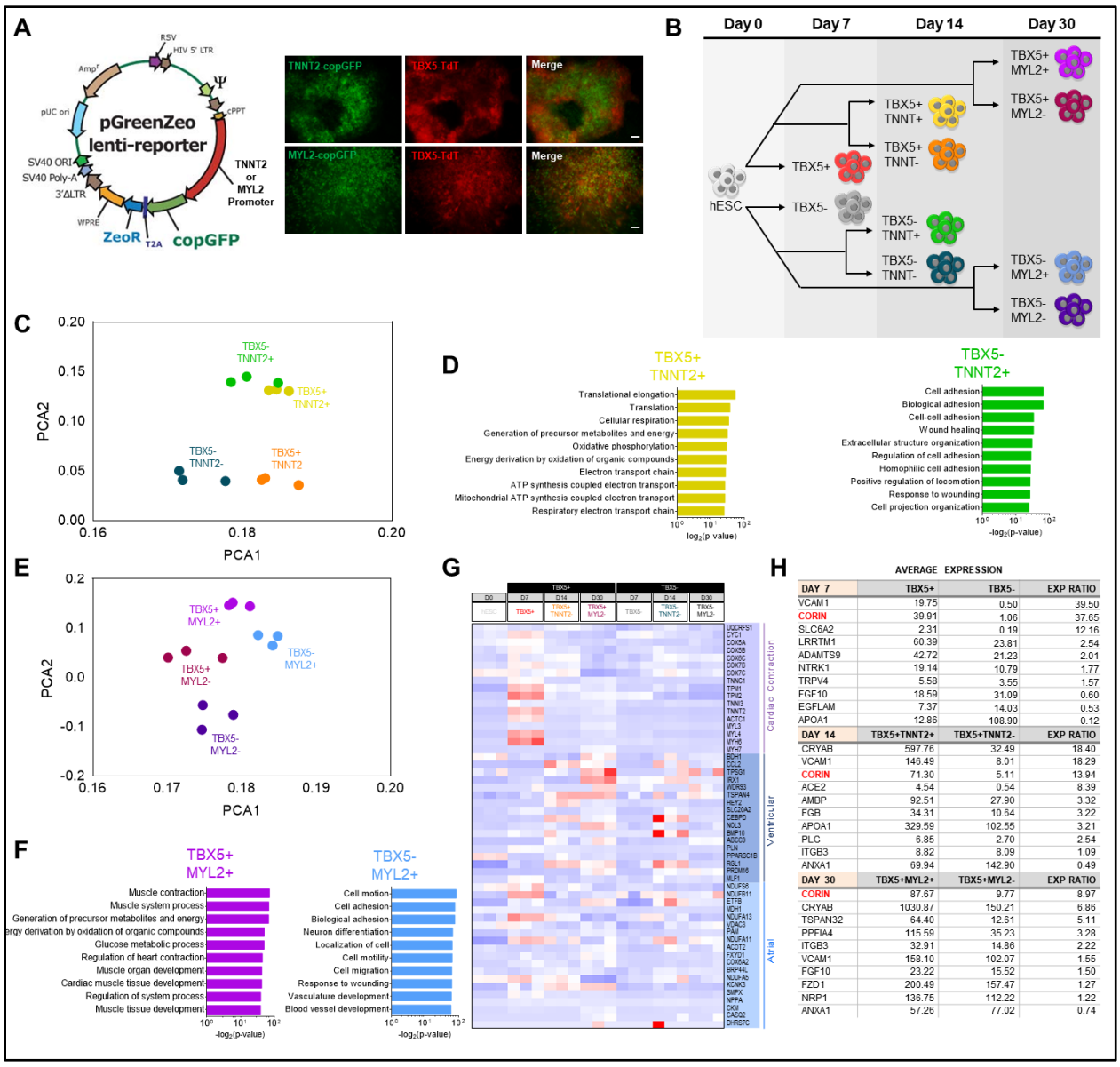
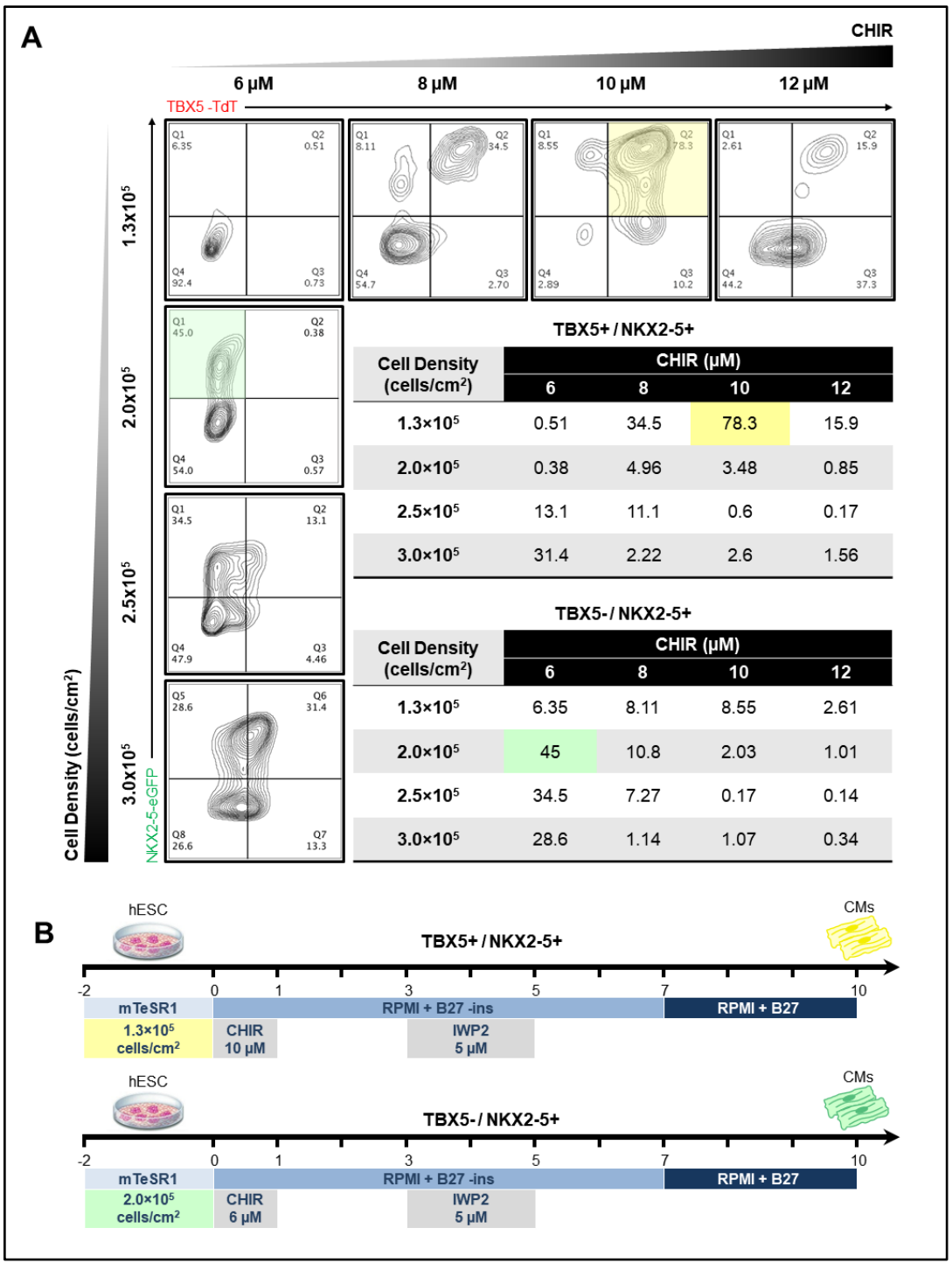


Figure 7. Transcriptional profiling of TBX5+ and TBX5- populations at different developmental stages. (A) Schematic of lentiviral constructs used to generate *TBX5*^{Tdtomato}/*TNNT2*^{copGFP} and *TBX5*^{Tdtomato}/*MYL2*^{copGFP} reporter lines and fluorescence microscopy images of hESC derived cardiomyocytes at day 14 (TNNT) and day 30 (MYL2) (Scale bars = 100µm). (B) Schematic showing all TBX5+ and TBX5- derived populations at varying timepoints of differentiation. (C) PCA analysis of TBX5+ and TBX5- cardiac and noncardiac populations at day 14. (D) GO analysis of the top differentially expressed genes between TBX5+TNNT2+ and TBX5-TNNT2+ populations at day 14 of cardiac differentiation. (E) PCA analysis of TBX5+ and TBX5- cardiac and noncardiac populations at day 30. (F) GO analysis of the top differentially expressed genes between TBX5+MYL2+ and TBX5-MYL2+ populations at day 30 of cardiac differentiation. (G) Heatmap comparing cardiac-related gene expression inTBX5+ and TBX5- derived non-cardiac populations. (H) Ranking of the top 10 expression ratios of cell surface genes between TBX5+ and TBX5- at day 7 and TBX5+ cardiac and non-cardiac populations at day 14 and 30. Scale bars = 100µm.

Generation and characterization of hESC-derived FHF-, SHF-, and nodal-like cardiomyocytes

Our *in vitro* and *in vivo* data suggest that both TBX5^{+} and to a lesser extent TBX5^{-} cells can give rise to CMs under cardiac differentiation conditions. RNA-sequencing analysis revealed that when used in conjunction with the mature CM markers TNNT2 or MYL2, TBX5^{+} cells can generate FHF-like CMs while the progeny of TBX5^{-} cells represent SHF-like CMs. To further explore the generation of heart field-specific CMs, we created a HES3-*TBX5*$^{TdTomato/W}$/*NKX2-5*$^{eGFP/W}$ double reporter line. We used the same TALEN targeting strategy as described above (Figure 2A) to insert the TdT gene into exon 9 of *TBX5* via a T2A peptide in a previously characterized HES3-*NKX2-5*$^{eGFP/W}$ hESC line [21]. Since NKX2-5 is expressed in nearly all myocytes of the developing heart (excluding the conduction system), we hypothesized that differentiation of this double reporter line would facilitate isolation of three distinct cardiac-related populations: 1) those that co-express TBX5 and NKX2-5 may represent FHF-like CMs (TBX5^{+}NKX2-5^{+}), 2) those that only express NKX2-5 may be SHF-like CMs (TBX5^{-}NKX2-5^{+}), and 3) cells that only express TBX5 may represent myocytes that make up the conduction system (TBX5^{+}NKX2-5^{-}). We optimized the differentiation conditions via altering seeding density and concentration of the GSK-3β inhibitor CHIR99021 to efficiently derive either double positive (TBX5^{+}NKX2-5^{+}) or single positive (TBX5^{-}NKX2-5^{+} or TBX5^{+}NKX2-5^{-}) populations (Figure 8A and 9A, colored boxes). Using a plating density of 1.3x10^{5} cells/cm^{2} with 10μM CHIR99021 or 2.0x10^{5} cells/cm^{2} with 6μM CHIR99021 in the differentiation protocol (Figure 8B) we were able to efficiently generate TBX5^{+}NKX2-5^{+} (76.7 +/- 1.3%) and TBX5^{-}NKX2-5^{+} cells (41.5 +/- 1.8%) respectively, as indicated by fluorescence microscopy, spontaneous beating,

and flow cytometry analysis at day 10 of differentiation (Figure 10A-D). When we modified our differentiation condition to $3.0x10^5$ cells/cm^2 with 12μM CHIR99021, we consistently observed generation of greater than 80% $TBX5^+NKX2\text{-}5^-$ cells (Figure 9A-C). Recent studies have reported that the sinoatrial node (SAN) originates from $NKX2\text{-}5^-$ progenitors in the right sinus horn, distinguishing them from ventricular and atrial myocytes that express NKX2-5 [56].



TBX5+ / NKX2-5+

Cell Density (cells/cm²)	CHIR (μM) 6	8	10	12
1.3×10⁵	0.51	34.5	78.3	15.9
2.0×10⁵	0.38	4.96	3.48	0.85
2.5×10⁵	13.1	11.1	0.6	0.17
3.0×10⁵	31.4	2.22	2.6	1.56

TBX5- / NKX2-5+

Cell Density (cells/cm²)	CHIR (μM) 6	8	10	12
1.3×10⁵	6.35	8.11	8.55	2.61
2.0×10⁵	45	10.8	2.03	1.01
2.5×10⁵	34.5	7.27	0.17	0.14
3.0×10⁵	28.6	1.14	1.07	0.34

Figure 8. Quantitative analysis of TBX5+NKX2-5+ and TBX5-NKX2-5+ populations via small-molecule modulation of Wnt signaling and cell seeding densities. (A) Representative flow cytometric analysis at day 10 of cardiac differentiation showing the expression of TBX5 and NKX2-5. Left column represents varying seeding densities and top row shows varying concentrations of CHIR99021. Inset table shows results from all combinations of seeding densities and CHIR concentrations. Yellow and green highlighted boxes represent the maximum percentage of TBX5+NKX2-5+ and TBX5-NKX2-5+ cells obtained, respectively. (B) Schematic representation of the optimized protocols for differentiation of TBX5+NKX2-5+ and TBX5-NKX2-5+ populations.

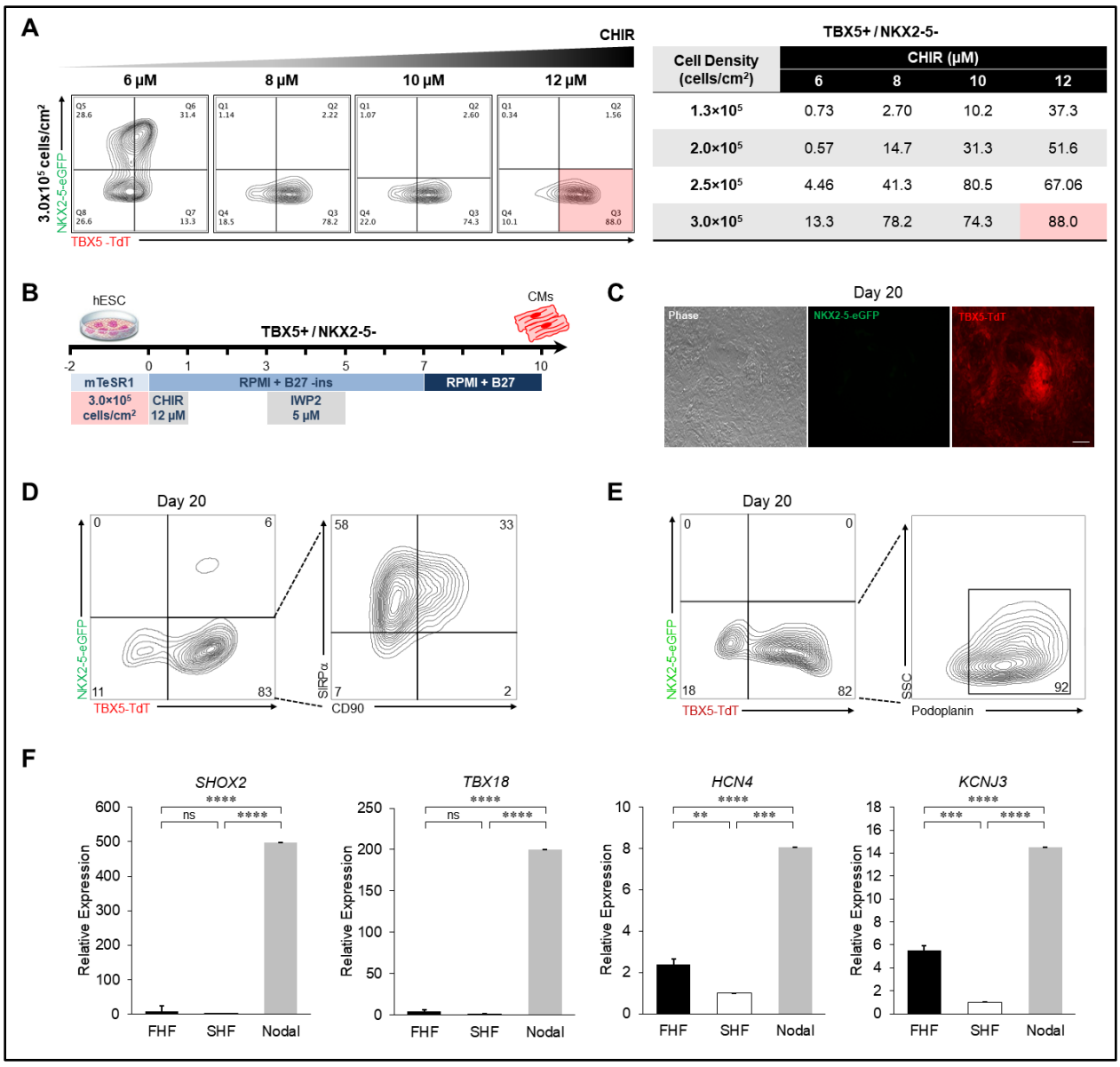


Figure 9. Optimization and characterization of hESC derived nodal-like cardiomyocytes. (A) Representative flow cytometric analysis at day 20 of cardiac differentiation to optimize the TBX5+NKX2-5- population by varying CHIR99021 concentration and seeding density. Table shows results from all combinations of seeding densities and CHIR concentrations. Red highlighted boxes represent the maximum percentage of TBX5+NKX2-5- cells obtained. (B) Schematic representation of the optimized protocol for differentiation of TBX5+ population. (C) Fluorescence microscopy of TBX5+NKX2-5- (nodal-like) optimized differentiation at day 20 (Scale bar = 100μm). (D) Flow cytometry analysis of TBX5+NKX2-5- cells for expression of SIRPα and CD90. (E) Flow cytometry analysis of TBX5+NKX2-5- cells for expression of Podoplanin. (F) Quantitative expression analysis of day 20 sorted nodal-like cardiomyocytes compared to day 10 sorted FHF- and SHF-like cardiomyocytes for nodal specific genes (n=3, Error bars represent SEM).

It has also been suggested that SAN-like pacemaker cells can be isolated from differentiating hESCs as NKX2-5^{-} CMs that express SAN lineage markers [57]. These findings prompted us to examine whether the TBX5^{+}NKX2-5^{-} cells may be enriched for SAN-like pacemaker cells. Flow cytometry analysis at day 20 showed that approximately 60% of the TBX5^{+}NKX2-5^{-} population was SIRPα^{+}CD90^{-}, indicative of a nodal cell phenotype (Figure 9D). The majority of these cells (92%) also expressed Podoplanin, a marker for cardiac pacemaker cells [58], confirming their nodal identity (Figure 9E).

To further characterize the identity of TBX5^{+}NKX2-5^{+}, TBX5^{-}NKX2-5^{+}, and TBX5^{+}NKX2-5^{-} as FHF-, SHF-, and nodal-like cells, respectively, we FACS isolated each population (from separate differentiation cultures) then performed RT-qPCR. Gene expression analysis revealed that the TBX5^{+}NKX2-5^{+} cells were enriched for the FHF markers, *TBX5* and *HAND1*, and CM marker *NKX2-5* while showing lower expression of multiple SHF markers (Figure 10E and 11A). On the other hand, the TBX5^{-}NKX2-5^{+} cells showed increased expression of SHF markers, and *NKX2-5* and lower expression of *TBX5* and *HAND1* (Figure 10E and 11A). Similar differences in expression pattern were observed in human fetal (gestation age 18 weeks) LV and RV samples (Figure 10F and 11B). The TBX5^{+}NKX2-5^{-}SIRPα^{+}CD90^{-} cell population showed high expression of *TBX5* and nodal cell markers, such as *SHOX2*, *TBX18*, *HCN4* and *KCNJ3* with low expression of *ISL1* and *NKX2-5* (Figure 10E and 11F). Immunocytochemistry analysis of TBX5^{+}NKX2-5^{+} and TBX5^{-}NKX2-5^{+} isolated cells supported the gene expression results and confirmed their CM phenotype with positive co-staining of TNNT2 and FHF or SHF markers while showing the absence of the nodal cell marker, SHOX2 (Figure 10G and 10H). Together these results confirm that the *TBX5*$^{TdTomato/W}$/*NKX2-5*$^{eGFP/W}$ double

reporter hESC line enables the prospective isolation of heart field-specific CMs and nodal cells from differentiating hESCs.

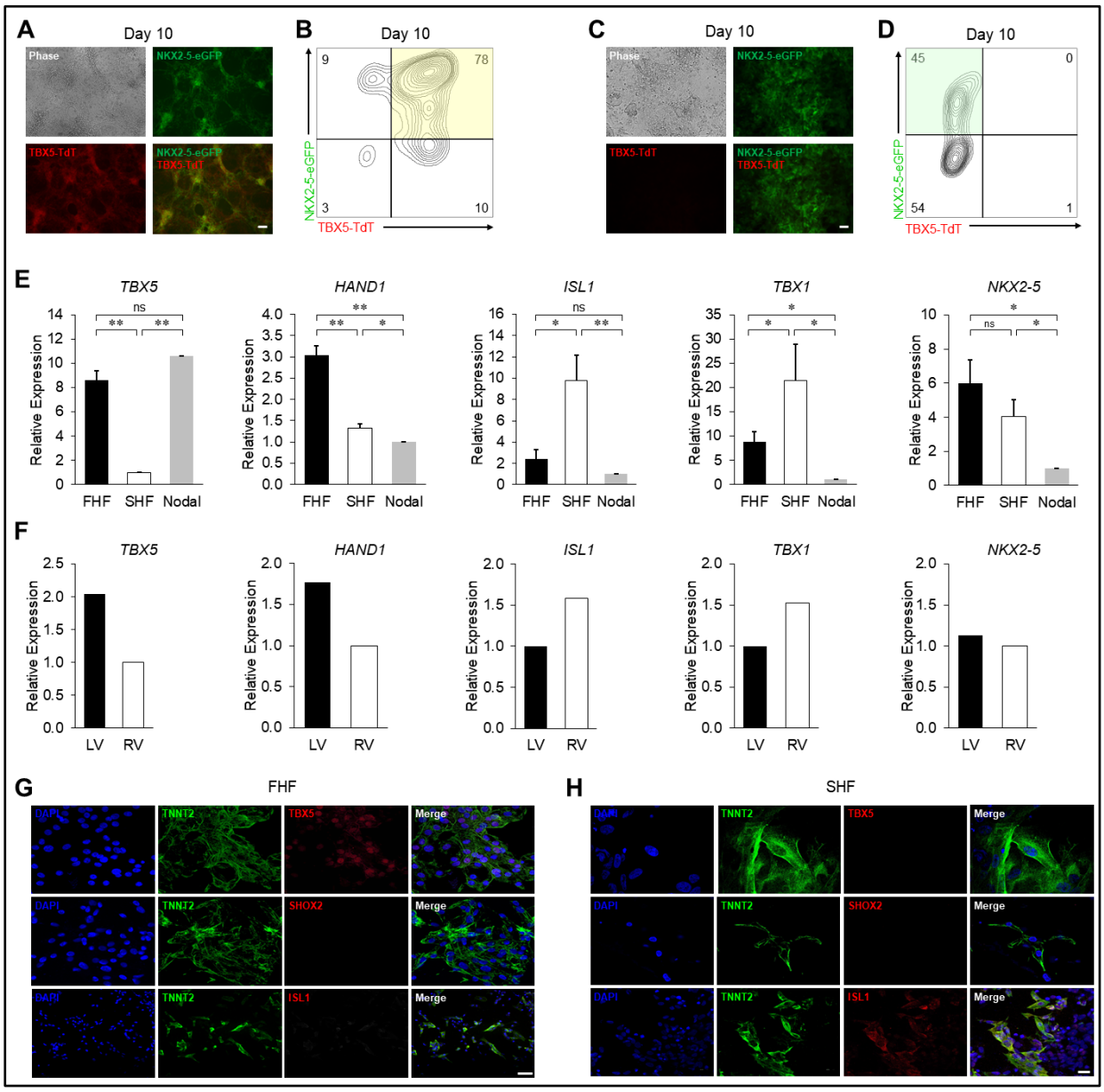


Figure 10. Characterization of hESC-derived FHF- and SHF-like cardiomyocytes. (A) Fluorescence microscopy and (B) representative flow cytometry plot of TBX5+NKX2-5+ (FHF) optimized differentiation at day 10. (C) Fluorescence microscopy and (D) representative flow cytometry plot of TBX5-NKX2-5+ (SHF) optimized differentiation at day 10. (E) Quantitative expression analysis of day 10 sorted FHF-, SHF- and nodal-like cardiomyocytes (n=3, Error bars represent SEM) and (F) isolated left and right ventricles (LV and RV, respectively) of human fetal heart for FHF (*TBX5* and *HAND1*), SHF (*TBX1* and *ISL1*) and *NKX2-5* genes (n=1). (G and H) Immunocytochemistry analysis of day 10 sorted FHF- and SHF-like cardiomyocytes for specific markers of FHF (TBX5), SHF (ISL1) and nodal cells (SHOX2).

Presence of cardiomyocytes is indicated by expression of TNNT2 and all nuclei were visualized with DAPI (G: Scale bar = 50µm, H: Scale bar = 20µm).

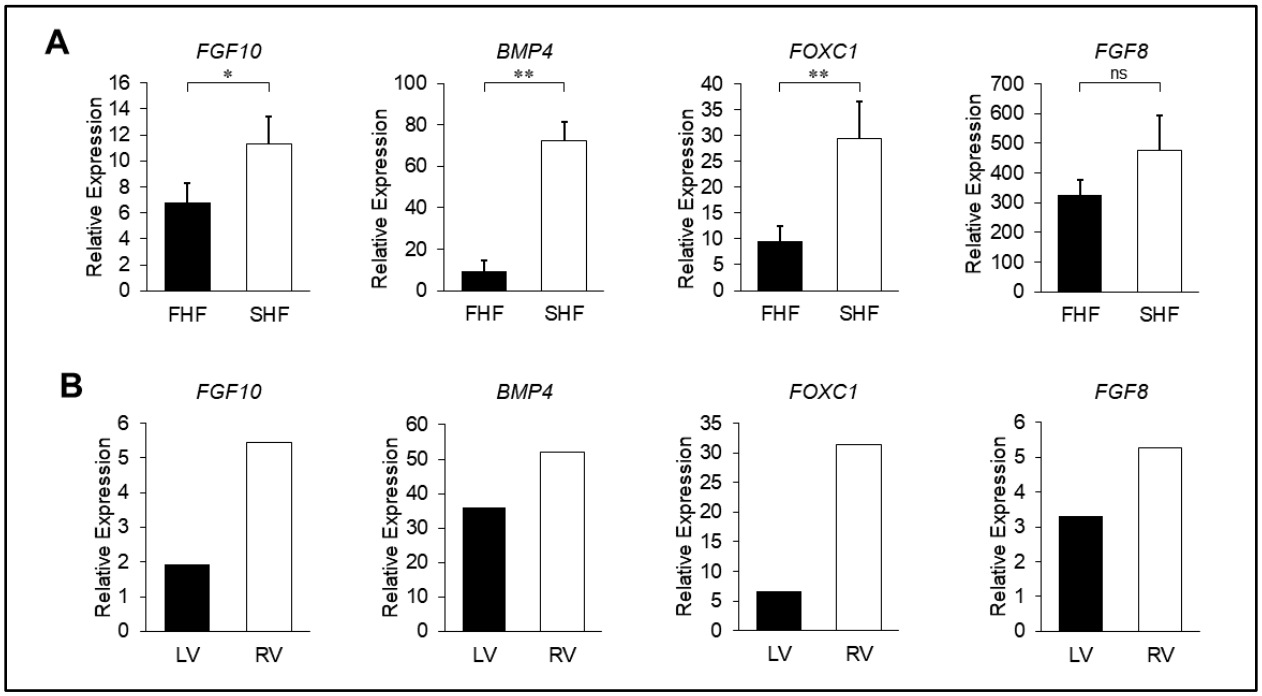


Figure 11. Characterization of hESC-derived FHF- and SHF-like cardiomyocytes. (A) Quantitative expression analysis of day 10 sorted FHF- and SHF-like cardiomyocytes (n=3, Error bars represent SEM) and (B) isolated left and right ventricles (LV and RV, respectively) of human fetal heart for additional SHF genes (n=1).

hESC-derived FHF- and SHF-like cardiomyocytes form functionally coupled monolayers and exhibit appropriate pharmacological responses

The myocardium forms a functional syncytium with individual CMs tightly coupled by adhesion junctions and connexins to provide mechanical and electrical connections. We sought to determine whether a 2-dimensional culture of hESC-derived first and second heart field-like CMs exhibit electromechanical connectivity by performing optical mapping. Day 10 FHF- and SHF-like CMs were isolated and re-plated to form monolayers. After 3 days of culture, robust spontaneous contractions were observed throughout the monolayers, which were then labeled with a voltage-sensitive dye. Optical

mapping signals demonstrated that the FHF- and SHF-like CM monolayers were spontaneously active (Figure 12A and 13A). Action potential propagation was recorded from both spontaneous and pace-induced cells in FHF-like CMs. Spontaneous activation was present (cycle length=3 seconds) at baseline (Figure 12A) while the higher pacing rates (0.5 Hz) in FHF-like CMs showed the emergence of 1:1 synchronous capture of the entire monolayer (Figure 12B). Isochronal activation map derived from optical mapping of FHF-like CM monolayers showed uniform action potential (AP) propagation throughout the monolayer with no block or any changes in conduction velocity (Figure 12C). In SHF-like CM cultures, spontaneous AP propagations were recorded from islands of cells (Figure 13A) in which APs propagated independently, but uniformly, in each island. To further analyze the identity of the CMs we performed single cell patch clamp, which revealed isolated hESC-derived CMs from both FHF and SHF differentiations had typical atrial and ventricular APs with fast upstroke velocities (> 30 V/s) and long AP duration (APD90>400ms) (Figure 13B).

We next examined the chronotropic responses of FHF- and SHF-like CMs to β-adrenergic drugs *in vitro*. Isoproterenol (1 μM), a β-adrenoreceptor agonist, enhanced the contraction rates of both FHF- and SHF-like CMs, whereas addition of Propranolol (Prop) (10 μM), a β-adrenoreceptor antagonist resulted in recovery to the baseline contraction rate (Figure 12D and 13C). Together, the functional analysis of FHF- and SHF-like CMs reflect the maturation, coordination of channel activities, presence of functional gap junctions, and the ability to respond to pharmacological agents in hESC- derived CMs.

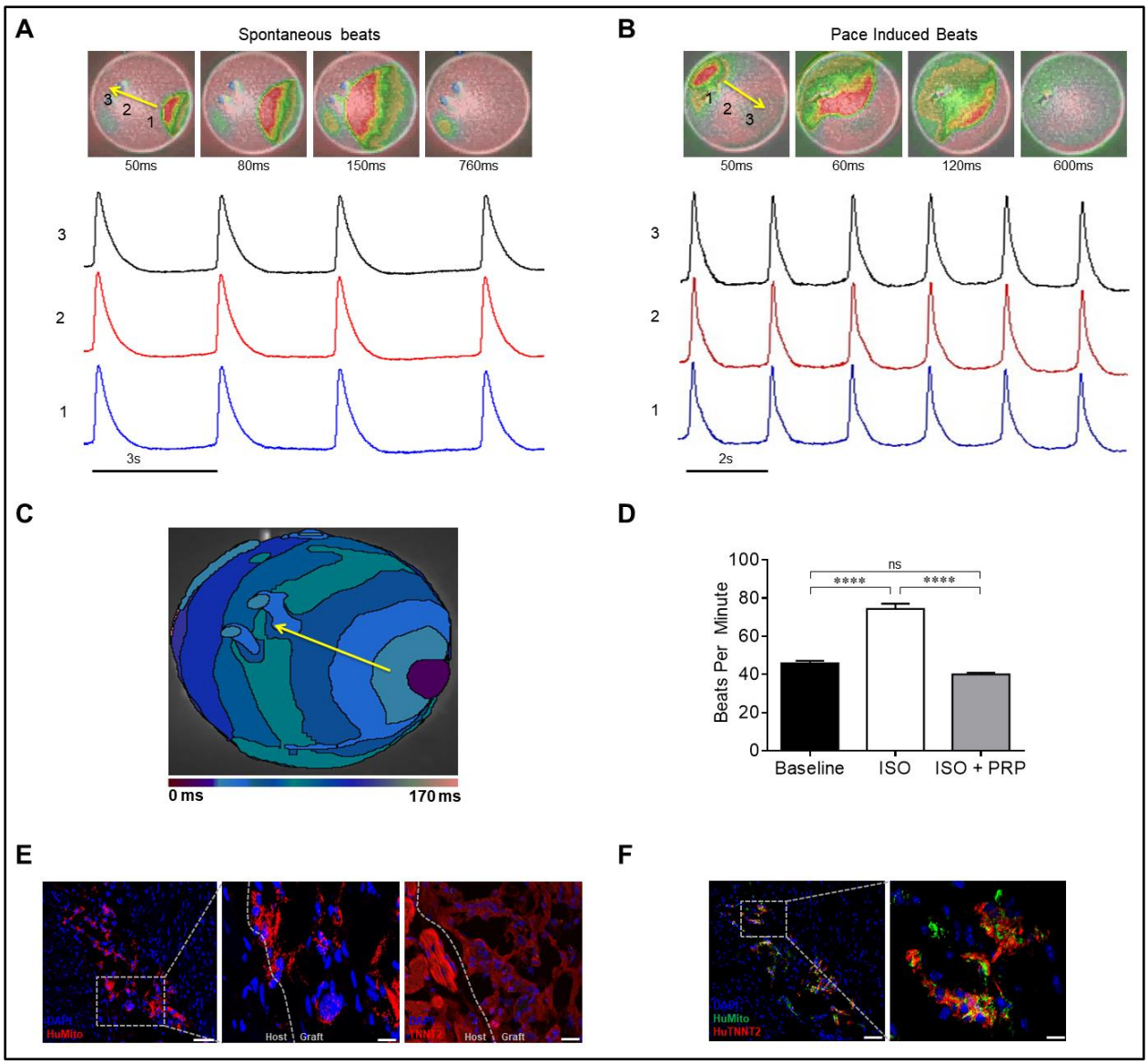


Figure 12. Functional analyses and transplantation of FHF-like cardiomyocytes. Snapshots of optical mapping show (A) spontaneous AP and (B) pace-induced AP which propagates across monolayer of hESC-derived cardiomyocytes (TBX5+NKX2-5+). The yellow arrow denotes direction of AP propagation. Numbers denote the area where APs were recorded. Red and Pink represent depolarization and repolarization phases, respectively. (C) Activation map showing uniform action potential propagation derived from optical mapping. (D) Beating rates of FHF-like cardiomyocytes following application of β–adrenergic agonist (isoproterenol, ISO, 1μM) and antagonist (propranolol, PRP, 10μM) (n=10, Error bars represent SEM). (E) Immunohistochemical analysis of engrafted FHF-like cardiomyocytes stained with a human specific mitochondria antibody (left) (Scale bar = 100μm), high magnification image of the boxed graft area (middle) and pan-species TNNT2 antibody staining of adjacent section (right) (Scale bars = 20μm). All cells were visualized using DAPI and dashed line denotes host/graft border. (F) Co-staining of mouse heart sections with human specific mitochondria and TNNT2 antibodies (Scale bar = 100μm). Higher magnification image of the boxed graft area is shown on the right (Scale bar = 20μm). All cells were visualized with DAPI.

Transplantation of FHF- and SHF-like cardiomyocytes into the injured mouse myocardium

To evaluate the *in vivo* survival and maturation of FHF- and SHF-like CMs, each fraction was sorted on day 10 of differentiation and cultured for an additional 3 days before transplantation into the left ventricular peri-infarct area of NSG mice seven days post-I/R injury. Human specific mitochondrial staining showed numerous clusters of FHF- and SHF-like CMs varying in size (Figure 12E and 13D). The transplanted cells exhibited expression of TNNT2 as expected, however, they showed less sarcomeric organization than the host CMs. Further confirmation of the graft area was carried out by co-staining using human specific antibodies to both mitochondria and TNNT2. Interestingly, detailed confocal microscopy revealed partially organized structural proteins in small areas of the FHF- and SHF-like CM grafts depicted by TNNT2 staining (Figure 12F and 13E). Together, these results show survival of hESC-derived FHF- and SHF-like CMs with a more organized ultra-structure in the graft area of injured mouse hearts.

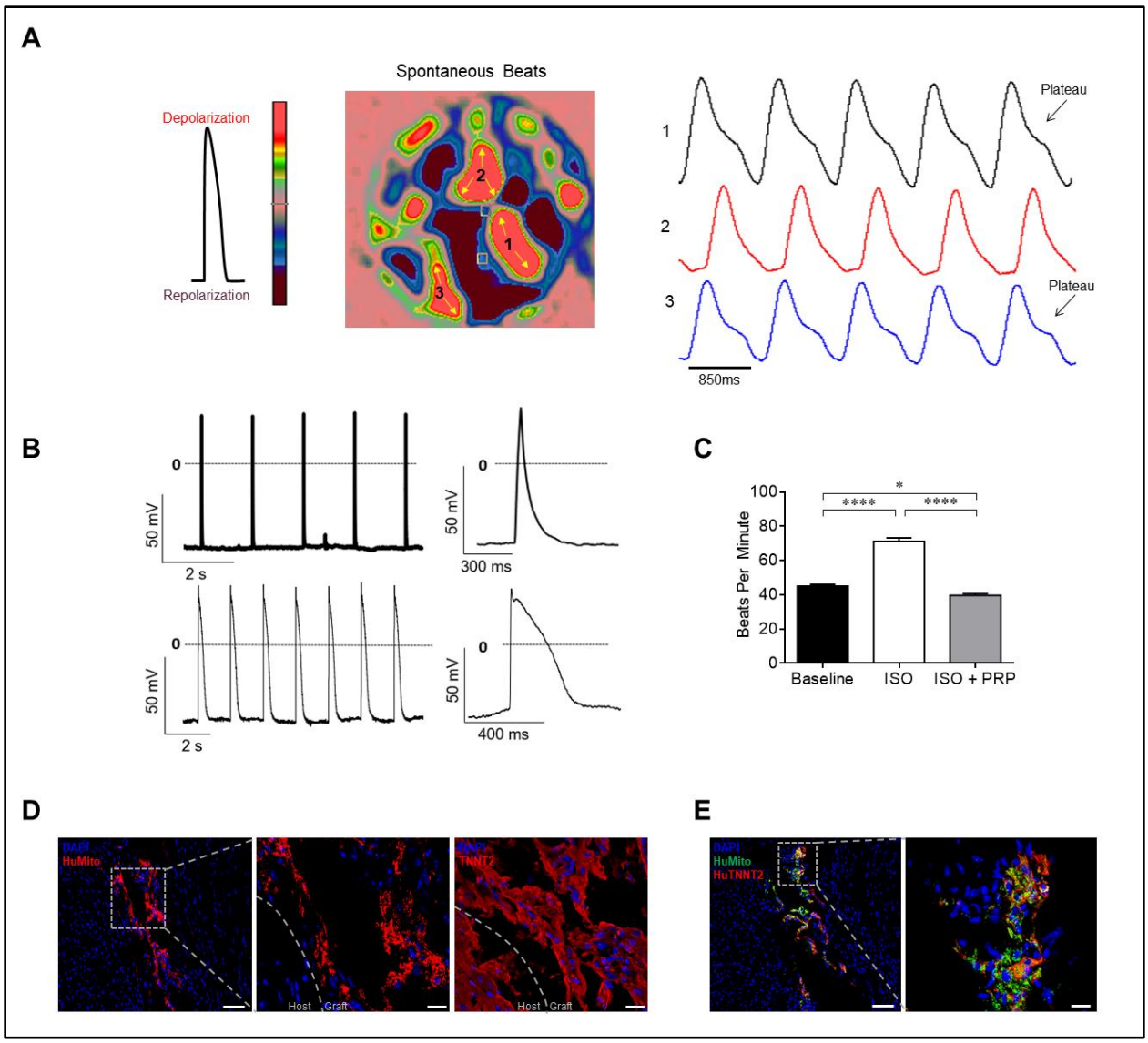


Figure 13. Functional analyses and transplantation of SHF-like cardiomyocytes. (A) Snapshot of optical mapping show spontaneous APs which propagates across islands of hESC-derived cardiomyocytes (TBX5-NKX2-5+). The yellow arrows denote direction of AP propagations. Numbers denote the islands of the cells from which APs were recorded. Red and Purple represent depolarization and repolarization phases, respectively. (B) Representative patch-clamp recordings of spontaneous APs of individual atrial and ventricular-like cardiomyocytes (n=10). (C) Beating rates of SHF-like cardiomyocytes following application of β–adrenergic agonist (isoproterenol, ISO, 1μM) and antagonist (propranolol, PRP, 10μM) (n=10, Error bars represent SEM). (D) Immunohistochemical analysis of engrafted SHF-like cardiomyocytes stained with a human specific mitochondria antibody (left) (Scale bar = 100μm), high magnification image of the boxed graft area (middle) and pan-species TNNT2 antibody staining of adjacent section (right) (Scale bar = 20μm). All cells were visualized using DAPI and dashed line denotes host/graft border. (E) Co-staining of mouse heart sections with human specific mitochondria and TNNT2 antibodies (Scale bar = 100μm). Higher magnification image of the boxed graft area is shown on the right (Scale bar = 20μm). All cells were visualized with DAPI.

CORIN is a surface marker for the isolation of FHF-like cardiomyocytes

Finally, we sought to isolate hESC-derived CMs from an independent non-genetically modified cell line. Our RNA-sequencing had revealed CORIN as a potential novel surface marker that is differentially expressed on CMs at both early and late time points of differentiation (Fig 5F and 7H). Time course analysis demonstrated that CORIN is expressed from day 7 and maintained until day 30 of cardiac differentiation (Figure 14A). Post-day 15 over 68% of cells maintained CORIN expression. Furthermore, CORIN is expressed on CMs from intact human heart tissues using immunohistochemistry on human left ventricle heart sections, which showed diffuse but selective staining of CMs (Figure 14B). These data support the notion that CORIN expression is maintained in CMs throughout development and into adult life.

To determine if CORIN can be used as a surface marker for specific isolation of CMs, we performed flow cytometry analysis for $CORIN^+$ and $CORIN^-$ cells after 10 days of FHF or SHF optimized differentiations. If the differentiation was biased towards FHF-like CMs, analysis revealed that 82% of the $CORIN^+$ population were FHF-like CMs ($TBX5^+NKX2\text{-}5^+$) (Figure 15A top) indicating the ability of CORIN to isolate FHF-like CMs. Additionally, selection for CORIN further enhanced the purity of FHF-like CMs when compared to an efficient FHF differentiation without CORIN selection, which contained ~30% contaminant cells (Figure 15A, bottom). A similar reduction in contaminant cells was also observed in a non-optimized differentiation, suggesting the specificity of CORIN to separate FHF-like CMs from other contaminant cells. Interestingly, when the

differentiation was biased towards SHF-like CMs, there were no CORIN$^+$ cells present in the SHF population (Figure 15B). These results indicate that CORIN represents a selective surface marker that can be used to isolate and enrich FHF-like CMs from differentiating hESCs. To further asses the ability of CORIN to specifically distinguish FHF and SHF CMs, left and right ventricles of human fetal hearts (gestation age 18 weeks) were isolated and analyzed for *CORIN* gene expression along with hESC-derived FHF- and SHF-like CMs. *CORIN* was expressed at a higher level in both the left ventricle and FHF-like CMs (Figure 14C), again validating its specificity for FHF CMs.

To confirm the FHF-like CM identity of CORIN$^+$ day 10 isolated cells, we performed RT-qPCR and immunocytochemistry (Figure 15C and 15D). As expected, CORIN$^+$ cells exhibited high expression of *TBX5*, but low expression of *ISL1* and *SHOX2* suggesting that these cells have a FHF-like CM phenotype. Both flow cytometry (Figure 15E) and immunocytochemistry (Figure 15F) for TNNT2 revealed that the majority (90%) of CORIN$^+$ cells were indeed TNNT2$^+$ CMs and displayed robust spontaneous contractions. We performed optical mapping to assess their electrophysiological properties. Action potential propagation was recorded from spontaneous beating cells in CORIN$^+$ CMs (Figure 15G). Optical mapping demonstrated that CORIN$^+$ cells have similar electrophysiological characteristics as the FHF-like CMs isolated from the double reporter hESC line (Figure 15G and 12A).

To evaluate the *in vivo* survival and maturation of CORIN$^+$ CMs, cells were transplanted into the left ventricular peri-infarct area of NSG mice seven days post-I/R injury. Human specific mitochondrial staining showed clusters of transplanted cells that

exhibited expression of TNNT2 as expected, however, they showed less sarcomeric organization than the host CMs. Co-staining using human specific antibodies to both mitochondria and TNNT2 revealed partially organized structural proteins in small areas of the graft (Figure 15H). These results show survival of CORIN$^+$ CMs in the graft area of injured mouse hearts similar to transplanted hESC-derived FHF-like CMs.

Taken together, our studies provide several independent lines of evidence that CORIN is a novel surface marker that can be used to isolate FHF-like CMs from a mixture of differentiating hESCs that may include SHF- and nodal-like CMs. CORIN$^+$ isolated CMs show similar characteristics to their FHF-like counterparts isolated using a double transgenic reporter line, thereby eliminating the need of genetically modified lines for isolation of heart field-specific CMs.

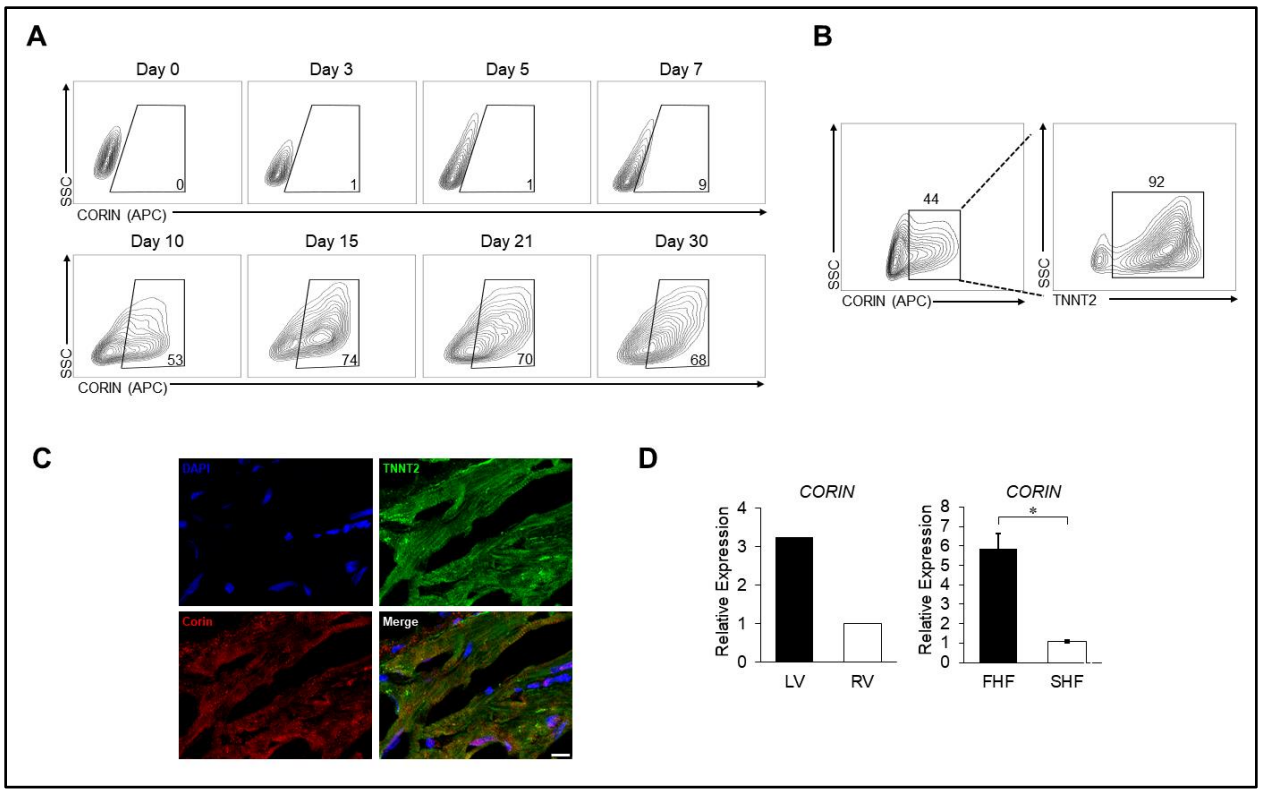


Figure 14. CORIN is a marker of FHF cardiomyocytes. (A) Flow cytometry time course analysis of CORIN expression in differentiating hESCs. (B) Flow cytometry analysis for TNNT2 in CORIN+ cells at day 10 of differentiation. (C) Immunohistochemistry analysis of CORIN expression in adult human left ventricular tissue (Scale bar = 20μm). (D) Quantitative PCR analysis of isolated left and right ventricles (LV and RV, respectively) of human fetal heart (n=1) and day 10 isolated FHF- and SHF-like cardiomyocytes for expression of *CORIN* (n=3, error bars represent SEM).

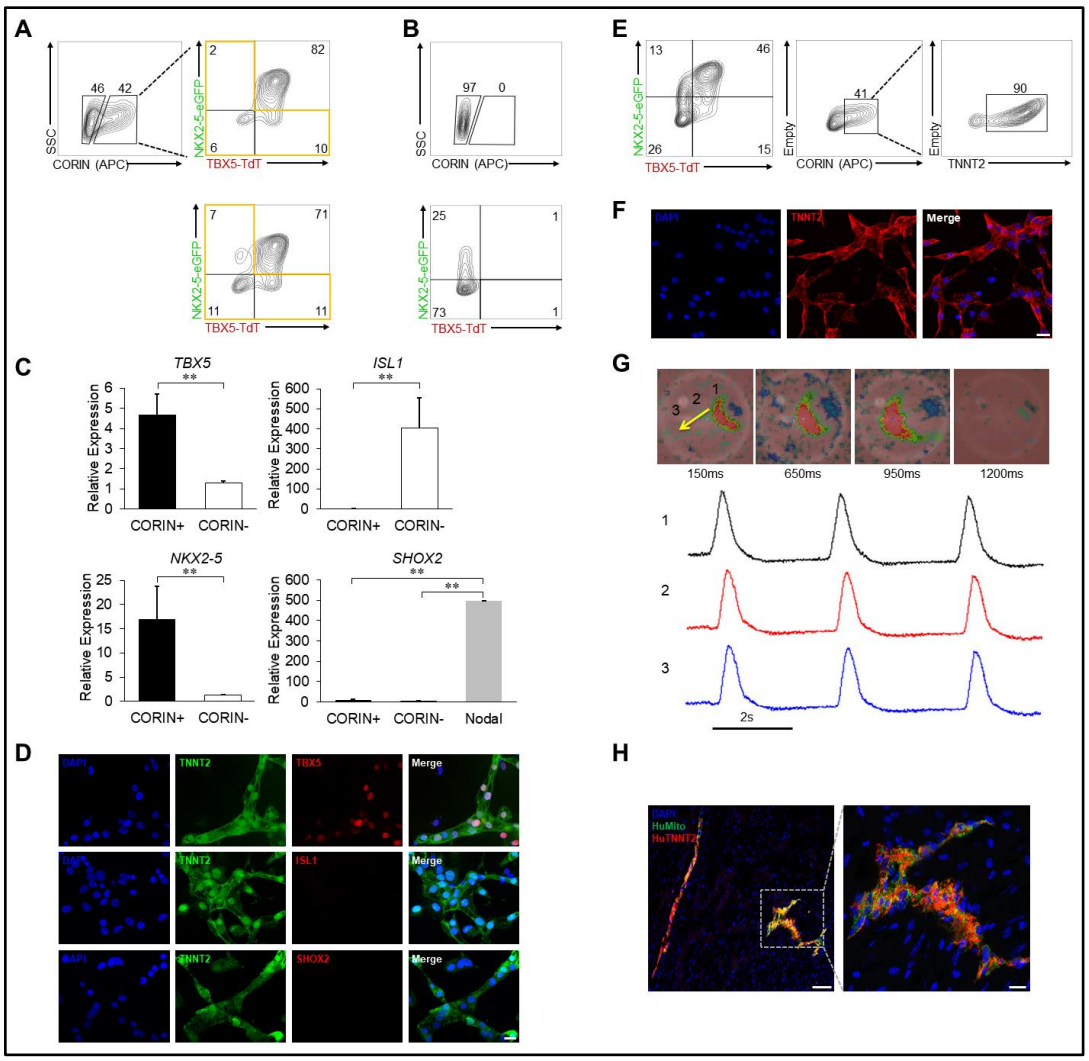


Figure 15. Isolation and enrichment of FHF-like cardiomyocytes using the surface marker CORIN. (A) Flow cytometric analysis of CORIN+ cells for FHF-like cardiomyocytes from an efficient FHF differentiation (top). Representative flow cytometry plot of same efficient FHF differentiation without CORIN selection (bottom). Contaminate (non-FHF) cells are outlined in orange box. (B) Flow cytometric analysis of CORIN+ cells from a SHF differentiation (top). Flow cytometric analysis of the same SHF differentiation shown on top based on TBX5 and NKX2-5 (bottom). (C) Quantitative PCR analysis of FHF (*TBX5*), SHF (*ISL1*), *NKX2-5* and nodal (*SHOX2)* genes in CORIN+ and CORIN- day 10 sorted cells. *SHOX2* expression was compared to day 20 sorted nodal-like cells (n=3, Error bars represent SEM). (D) Immunocytochemistry of day 10 sorted CORIN+ cardiomyocytes for markers of FHF (TBX5), SHF (ISL1) and nodal cells (SHOX2). Cardiomyocytes are shown by TNNT2 staining and all nuclei were shown by DAPI (Scale bar = 20µm). (E) Flow cytometry analysis of TNNT2 in CORIN+ cells (right) from a moderate FHF differentiation (left). (F) Immunocytochemistry of CORIN+ sorted cells for TNNT2. All nuclei are shown by DAPI (Scale bar = 20µm). (G) Snapshots of optical mapping show spontaneous AP propagates across monolayer of CORIN+ cardiomyocytes. The yellow arrow denotes direction of AP propagation. Numbers denote the area from which the APs were recorded. Red and Pink represent depolarization and repolarization phases, respectively. (H) Immunohistochemical analysis of engrafted CORIN+ cardiomyocytes in mouse heart sections co-stainined with human specific mitochondria and TNNT2 antibodies (Scale bar = 100µm). Higher magnification image of the boxed graft area is shown on the right (Scale bar = 20µm). All cells were visualized with DAPI.

Discussion

Despite recent advances in the treatment of heart disease, acute myocardial infarction (MI) is still associated with significant morbidly and mortality. Based on some promising preclinical studies there is increasing interest in the use of pluripotent stem cells to promote cardiac repair after acute MI [34, 59-63]. Previous attempts to generate hESC-derived CMs rely mainly on the observation of contracting cells in a dish, which not only include impurities such as non-myocyte cells, but also a mixture of CM subtypes including ventricular, atrial and pacemaker cells. The failure to isolate a pure population of chamber specific cardiomyocyte may compromise the safety of cell-based transplantation for future clinical applications. Furthermore, the delivery of a cell population contaminated by pacemaker cells could make the graft a focus for arrhythmias. This possibility may have contributed to the observed ventricular arrhythmias in previous reports where hESC/hiPSC-derived CMs were transplanted in non-human primates [34, 36]. Therefore, the safety and efficacy of cell-based therapies will be vastly improved with the delivery of a pure population of chamber specific CMs.

Following an acute MI, most commonly caused by blockage of the left coronary artery, there is extensive and irreversible loss of CMs with the damage primarily localized to the LV of the heart. The CMs of the LV are predominantly derived from FHF CPCs whereas CMs within the RV originate from the SHF during early cardiogenesis. To our knowledge no previous study has demonstrated the ability to purify FHF CMs from hESCs for preclinical transplantation studies. Previous studies have attempted to generate heart field-specific CMs using reporter lines that are considered specific for the SHF, however

these lines generated heterogeneous cell populations containing both cardiac and non-cardiac cell types [19, 37, 43].

In this study, we focused on TBX5, which is an important transcription factor in the development of the FHF. We generated a hESC reporter line to investigate the developmental potential of TBX5 CPCs to give rise to cardiac cell types *in vitro* and *in vivo*. Interestingly in contrast to our *in vitro* findings, *in vivo* transplantation studies revealed that both $TBX5^+$ and $TBX5^-$ CPCs predominantly become CMs, which implies that the *in vivo* environment may influence the plasticity of hESC-derived CPCs. Furthermore, we observed clear differential expression of SHF genes in the $TBX5^-$ populations using transcriptomic analyses, which supported our hypothesis that FHF and SHF CMs can be separated at this phase of differentiation, even within a two-dimensional culture environment that lacks the spatial information present in the embyro.

Since the expression of TBX5 is necessary for the development of a variety of myocytes, including the conduction system of the heart [64], our single reporter hESC line most likely generated a heterogeneous population of CMs including pacemaker cells. On the other hand, NKX2-5 is an essential transcription factor that regulates the formation of nearly all ventricular and atrial CMs, but not pacemaker cells [56]. In order to develop a hierarchy for hESC-derived CPC development, we constructed a double reporter line utilizing TBX5 and NKX2-5, which allowed us to not only isolate and enrich FHF- and SHF-like CMs, but also exclusively separate nodal-like cells. To investigate their electrophysiological properties and assess their ability to form a functional syncytium we performed optical mapping studies on the isolated FHF- and SHF-like CMs. As expected,

these studies led to the observation that our FHF- and SHF-like CMs are predominantly composed of atrial and ventricular CMs. They also displayed similar electrical wavefront activation establishing electrical coupling and regenerative AP propagation during spontaneous depolarizations and following external electrical pacing. The optical AP profiles were remarkably similar to those measured using patch-clamp. Conduction velocity (CV) in 2 dimensional cell layers were much slower than reported in human CMs [65]. While several factors could contribute to this slow CV (i.e. CM connectivity and gap junction densities), a major factor is likely the relatively positive resting diastolic membrane potential which promotes sodium channel inactivation [66]. Despite these observations, generation of functional ventricular and atrial CM subtypes as characterized by marker expression and electrophysiological analyses, would be of great interest for safe regenerative medicine, drug discovery, and human arrhythmia modeling.

Finally, to achieve full translational application of our findings, it is important to isolate FHF CMs without using genetically modified hESC lines. We discovered VCAM1, CRYAB, and CORIN as 3 potential surface markers when we compared their expression ratios at 3 differentiation time points (Figure 7H). VCAM1 has been previously described to mark early cardiovascular lineage commitment in differentiating hESC/hiPSC as well as mature CMs [21, 55, 67]. However, our expression ratio data from day 30 suggests that more robust surface markers for mature CMs may exist. Crystallin alpha B (CRYAB) is a member of the small heat shock protein 20 family whose function has been studied in cardiac hypertrophy and myopathies but has not been studied as a surface marker for CMs [68-71]. CRYAB was not ranked in the top 10 expression ratios at day 7 suggesting that it may be a more suitable marker for later stage CMs. CORIN is a serine protease

responsible for the conversion of pro-atrial natriuretic peptide (ANP) and pro-brain natriuretic peptide (BNP) to their active forms, ANP and BNP respectively, in CMs and has been suggested as a biomarker in hypertension, cardiac hypertrophy, and heart failure [72-77]. CORIN ranked in the top 3 expression ratios at all 3-time points, which prompted us to further pursue the ability of CORIN to isolate CMs from differentiating hESCs. Interestingly, CORIN is activated by the proprotein convertase subtilisin/kexin type 6 enzyme (PCSK6), which has been shown to be co-expressed with *TBX5* and involved in development of left and right asymmetry [78] suggesting a potential role for CORIN in early development. In support of these observations, we show that CORIN can be used as a surface marker to selectively isolate FHF-like CMs from differentiating hESCs throughout their development from a progenitor to a more mature state.

Our findings allow, for the first time, prospective isolation of FHF-like CMs from SHF-like CMs in differentiating hESCs. While their developmental pattern and epigenetic background can now be investigated *in vitro*, there is still considerable controversy as to whether there is a functional difference between a LV and RV CM, beyond the pressure and shear forces they are exposed to in fetal or adult hearts. Although left and right ventricular CMs emerge from different developmental origins, they express similar ion channels, structural proteins, and display similar electrical properties. Nonetheless, in congenital heart defects where the RV acts as a systemic ventricle, there is generally ventricular dysfunction over time, which suggests that there may be inherent differences between right and left ventricular CMs. It remains unknown whether transplantation of RV CMs in the damaged LV will demonstrate similar incompetency. Therefore, isolation of

pure FHF CMs will allow for elimination of possible incompetent RV CMs from the transplanted cells.

In summary, here we describe a practical approach for isolation and characterization of hESC-derived first and second heart field-like CMs as well as nodal-like cells. Access to enriched populations of CM subtypes is important for treating diseases that affect specific regions of the heart, such as myocardial infarction involving the left and right ventricles or chamber-specific congenital heart defects. Furthermore, heart field-specific CMs can be used for developmental studies, arrhythmia modeling, drug screening and pre-clinical regenerative therapy studies in which pure populations of cells are essential.

Tables

Probe		PCR probe size	Enzyme for Southern	Expected Southern Bands
5’ external probe				
TBX-5f1	5’-ATCCATCCATCCATCCATCCATCC-3’	582 bp	NsiI	Targeted: 3.3kb WT: 12.4kb
TBX-5r1	5’AACTAACCAGGTGCAGTTGTGCAC-3’			
3’ external probe				
TBX-3f1	5’AGGGTAAGATGTGAGGCTGAGTG-3’	650 bp	NdeI	Targeted: 3.1kb WT: 8.6kb
TBX-3r1	5’-GAGCTCAAGTTCACAAATCCTAG-3’			

Table 1. Southern blotting probes used in this study.

Gene	Forward 5’-3’	Reverse 5’-3’
TBX5	GACTTCCTACCAGAACCACAAG	GGGACCACGGGATATTCTTTAC
ISL1	ACCTTGGAAAGTACTGAGCG	GGTGTATCTGGAAGTTGAGAGG
NKX2-5	AGTGTGCGTCTGCCTTTC	GTTGTCCGCCTCTGTCTTC
ACTN2	CAAACCTGACCGGGGAAAAAT	CTGAATAGCAAAGCGAAGGATGA
TNNT2	AGCGGAAAAGTGGGAAGAG	TCCAAGTTATAGATGCTCTGCC
MYL2	CATGGCACCTAAGAAAGCAAAG	ATGAAGCCATCCCTGTTCTG
MEF2C	TGTAACACATCGACCTCCAAG	TGTTCAAGTTACCAGAGAC
GATA4	AGATGGGACGGGTCACTATC	CAGTTGGCACAGGAGAGG
MYL7	GACCCAGGCAGACAAGTTC	CCCTCATTCCTCTTTCTCGTC
CNN1	AACCATACACAGGTGCAGTC	GATGTTCCGCCCTTCTCTTAG
CD31	AACAGTGTTGACATGAAGAGCC	TGTAAAACAGCACGTCATCCTT

SHOX2	AGACCCACTATCCCGACG	GCCCCTATGAGAACACCTTTATG
TBX18	GTACATTCATCCAGACTCGCC	GCGGTTGGTATTTGTGCATAG
HCN4	TCGACTCGGAGGTCTACAAG	GGTCGTAGGTCATGTGGAAG
KCNJ3	GCCCCTTTTATGACCTATCCC	ACGATGACCCCAAAGAACTTC
HAND1	CCAAGGATGCACAGTCTGG	AGGAGGAAAACCTTCGTGCTG
TBX1	CGTGCAGCTAGAGATGAAGG	CATATAGTCGGCCATGGGATC
FGF10	GCCGTCAAAGCCATTAACAG	TCCTCTATCCTCTCCTTCAGC
BMP4	GCACTGGTCTTGAGTATCCTG	TGCTGAGGTTAAAGAGGAAACG
FOXC1	AGTAGCTGTCAAATGGCCTTC	TTAGTTCGGCTTTGAGGGTG
FGF8	GACCTACCAACTCTACAGCC	ACTCGAACTCTGCTTCCAAAG
CORIN	TCTGTTCTTGGGATTGTGGG	AGCGGCACTTGAAATGACTAG
GAPDH	ACATCGCTCAGACACCATG	TGTAGTTGAGGTCAATGAAGGG

Table 2. PCR Primers used in this study.

ANTIBODY	COMPANY	CATALOG #
Mouse monoclonal to Human SIRPα/β (clone SE5A5), PE/Cy7 conjugated	Biolegend	Cat.# 323807
Mouse monoclonal to Human VCAM1 (clone STA), APC conjugated	Biolegend	Cat.# 305809
Mouse monoclonal to Human TRA-1-81 (clone TRA-1-81), Alexa Fluor 594 conjugated	Biolegend	Cat.# 330712
Mouse monoclonal to Human CD90 (clone 5E10), APC conjugated	Biolegend	Cat.# 328113
Rat monoclonal to Human Podoplanin (clone NC-08), PE/Cy7 conjugated	Biolegend	Cat.# 337013
Mouse monoclonal to CORIN	Gift from Dr. Y. Ono, KAN Research Institute, JPN	N/A
Mouse monoclonal to TNNT2 (clone 13-11), BV421 conjugated	BD Biosciences	Cat.# 565618
Goat polyclonal to mouse IgG (clone Poly4053), APC conjugated	Biolegend	Cat.# 405308
Rabbit polyclonal to TNNT2	Abcam	Cat.# ab45932
Mouse monoclonal to TNNT2 (clone 13-11)	Thermo Fisher	Cat.# MA5-12960

Mouse monoclonal to TBX5 (clone A-6)	Santa Cruz	Cat.# sc-515536
Goat polyclonal to ISL1	R&D Systems	Cat.# AF1837
Mouse monoclonal to SHOX2	Abcam	Cat.# ab55740
Rabbit polyclonal to OCT4	EMD Millipore	Cat.# AB3209
Mouse monoclonal to NANOG (clone 7F7.1)	EMD Millipore	Cat.# MABD24
Mouse monoclonal to Human mitochondria (clone 113-1)	Abcam	Cat.# ab92824
Rabbit polyclonal to CD31	Abcam	Cat.# ab28364
Rabbit polyclonal to CNN1	Sigma	Cat.# HPA014263
Rabbit polyclonal to Cx43	Cell Signaling Technology	Cat.# 3512
Rabbit polyclonal to CORIN	Abcam	Cat.# ab209963
Goat anti-rabbit IgG (H+L), Alexa Fluor 488 conjugated	Thermo Fisher	Cat.# A-11008
Goat anti-mouse IgG (H+L), Alexa Fluor 488 conjugated	Thermo Fisher	Cat.# A-11001
Goat anti-rabbit IgG (H+L), Alexa Fluor 594 conjugated	Thermo Fisher	Cat.# A-11012
Goat anti-mouse IgG (H+L), Alexa Fluor 594 conjugated	Thermo Fisher	Cat.# A-11005
Rabbit anti-mouse IgG (H+L), Alexa Fluor 647 conjugated	Thermo Fisher	Cat.# A-21239
Donkey anti-goat IgG (H+L), Alexa Fluor 647 conjugated	Thermo Fisher	Cat.# A-21447

Table 3. Antibodies used in this study.

References

1. Garcia-Martinez V and Schoenwolf GC. Primitive-streak origin of the cardiovascular system in avian embryos. *Dev Biol*. 1993;159:706-19.

2. Yutzey KE and Bader D. Diversification of cardiomyogenic cell lineages during early heart development. *Circ Res*. 1995;77:216-9.

3. Jagla T, Bidet Y, Da Ponte JP, Dastugue B and Jagla K. Cross-repressive interactions of identity genes are essential for proper specification of cardiac and muscular fates in Drosophila. *Development*. 2002;129:1037-47.

4. Ward EJ and Skeath JB. Characterization of a novel subset of cardiac cells and their progenitors in the Drosophila embryo. *Development*. 2000;127:4959-69.

5. Frasch M. Induction of visceral and cardiac mesoderm by ectodermal Dpp in the early Drosophila embryo. *Nature*. 1995;374:464-7.

6. Park M, Wu X, Golden K, Axelrod JD and Bodmer R. The wingless signaling pathway is directly involved in Drosophila heart development. *Dev Biol*. 1996;177:104-16.

7. Rones MS, McLaughlin KA, Raffin M and Mercola M. Serrate and Notch specify cell fates in the heart field by suppressing cardiomyogenesis. *Development*. 2000;127:3865-76.

8. Riechmann V, Irion U, Wilson R, Grosskortenhaus R and Leptin M. Control of cell fates and segmentation in the Drosophila mesoderm. *Development*. 1997;124:2915-22.

9. Srivastava D. Making or breaking the heart: from lineage determination to morphogenesis. *Cell*. 2006;126:1037-48.

10. Icardo JM. Heart anatomy and developmental biology. *Experientia*. 1988;44:910-

11. Lohr JL and Yost HJ. Vertebrate model systems in the study of early heart development: Xenopus and zebrafish. *Am J Med Genet*. 2000;97:248-57.

12. Wagner M and Siddiqui MA. Signal transduction in early heart development (I): cardiogenic induction and heart tube formation. *Exp Biol Med (Maywood)*. 2007;232:852-65.

13. Buckingham M, Meilhac S and Zaffran S. Building the mammalian heart from two sources of myocardial cells. *Nat Rev Genet*. 2005;6:826-35.

14. Waldo KL, Kumiski DH, Wallis KT, Stadt HA, Hutson MR, Platt DH and Kirby ML. Conotruncal myocardium arises from a secondary heart field. *Development*. 2001;128:3179-88.

15. Anderson RH, Brown NA and Moorman AF. Development and structures of the venous pole of the heart. *Dev Dyn*. 2006;235:2-9.

16. Mummery CL, Zhang J, Ng ES, Elliott DA, Elefanty AG and Kamp TJ. Differentiation of human embryonic stem cells and induced pluripotent stem cells to cardiomyocytes: a methods overview. *Circ Res*. 2012;111:344-58.

17. Murry CE and Keller G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell*. 2008;132:661-80.

18. Passier R, Denning C and Mummery C. Cardiomyocytes from human embryonic stem cells. *Handb Exp Pharmacol*. 2006:101-22.

19. Bu L, Jiang X, Martin-Puig S, Caron L, Zhu S, Shao Y, Roberts DJ, Huang PL, Domian IJ and Chien KR. Human ISL1 heart progenitors generate diverse multipotent cardiovascular cell lineages. *Nature*. 2009;460:113-7.

20. Den Hartogh SC, Schreurs C, Monshouwer-Kloots JJ, Davis RP, Elliott DA, Mummery CL and Passier R. Dual reporter MESP1 mCherry/w-NKX2-5 eGFP/w hESCs enable studying early human cardiac differentiation. *Stem Cells*. 2015;33:56-67.

21. Elliott DA, Braam SR, Koutsis K, Ng ES, Jenny R, Lagerqvist EL, Biben C, Hatzistavrou T, Hirst CE, Yu QC, Skelton RJ, Ward-van Oostwaard D, Lim SM, Khammy O, Li X, Hawes SM, Davis RP, Goulburn AL, Passier R, Prall OW, Haynes JM, Pouton CW, Kaye DM, Mummery CL, Elefanty AG and Stanley EG. NKX2-5(eGFP/w) hESCs for isolation of human cardiac progenitors and cardiomyocytes. *Nat Methods*. 2011;8:1037-40.

22. Kitajima S, Miyagawa-Tomita S, Inoue T, Kanno J and Saga Y. Mesp1-nonexpressing cells contribute to the ventricular cardiac conduction system. *Dev Dyn*. 2006;235:395-402.

23. Kitajima S, Takagi A, Inoue T and Saga Y. MesP1 and MesP2 are essential for the development of cardiac mesoderm. *Development*. 2000;127:3215-26.

24. Lints TJ, Parsons LM, Hartley L, Lyons I and Harvey RP. Nkx-2.5: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants. *Development*. 1993;119:969.

25. Lyons I, Parsons LM, Hartley L, Li R, Andrews JE, Robb L and Harvey RP. Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. *Genes Dev*. 1995;9:1654-66.

26. Prall OW, Menon MK, Solloway MJ, Watanabe Y, Zaffran S, Bajolle F, Biben C, McBride JJ, Robertson BR, Chaulet H, Stennard FA, Wise N, Schaft D, Wolstein O, Furtado MB, Shiratori H, Chien KR, Hamada H, Black BL, Saga Y, Robertson EJ, Buckingham ME and Harvey RP. An Nkx2-5/Bmp2/Smad1 negative feedback loop controls heart progenitor specification and proliferation. *Cell*. 2007;128:947-59.

27. Saga Y, Kitajima S and Miyagawa-Tomita S. Mesp1 expression is the earliest sign of cardiovascular development. *Trends Cardiovasc Med*. 2000;10:345-52.

28. Saga Y, Miyagawa-Tomita S, Takagi A, Kitajima S, Miyazaki J and Inoue T. MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube. *Development*. 1999;126:3437-47.

29. Tanaka M, Wechsler SB, Lee IW, Yamasaki N, Lawitts JA and Izumo S. Complex modular cis-acting elements regulate expression of the cardiac specifying homeobox gene Csx/Nkx2.5. *Development*. 1999;126:1439-50.

30. Kelly RG. The second heart field. *Curr Top Dev Biol*. 2012;100:33-65.

31. Zaffran S and Kelly RG. New developments in the second heart field. *Differentiation*. 2012;84:17-24.

32. Dyer LA and Kirby ML. The role of secondary heart field in cardiac development. *Dev Biol*. 2009;336:137-44.

33. Caspi O, Huber I, Kehat I, Habib M, Arbel G, Gepstein A, Yankelson L, Aronson D, Beyar R and Gepstein L. Transplantation of human embryonic stem cell-derived cardiomyocytes improves myocardial performance in infarcted rat hearts. *J Am Coll Cardiol*. 2007;50:1884-93.

34. Chong JJ, Yang X, Don CW, Minami E, Liu YW, Weyers JJ, Mahoney WM, Van Biber B, Cook SM, Palpant NJ, Gantz JA, Fugate JA, Muskheli V, Gough GM, Vogel KW, Astley CA, Hotchkiss CE, Baldessari A, Pabon L, Reinecke H, Gill EA, Nelson V, Kiem HP, Laflamme MA and Murry CE. Human embryonic-stem-cell-derived cardiomyocytes regenerate non-human primate hearts. *Nature*. 2014;510:273-7.

35. Laflamme MA, Chen KY, Naumova AV, Muskheli V, Fugate JA, Dupras SK, Reinecke H, Xu C, Hassanipour M, Police S, O'Sullivan C, Collins L, Chen Y, Minami E, Gill EA, Ueno S, Yuan C, Gold J and Murry CE. Cardiomyocytes derived from human embryonic stem cells in pro-survival factors enhance function of infarcted rat hearts. *Nat Biotechnol*. 2007;25:1015-24.

36. Shiba Y, Gomibuchi T, Seto T, Wada Y, Ichimura H, Tanaka Y, Ogasawara T, Okada K, Shiba N, Sakamoto K, Ido D, Shiina T, Ohkura M, Nakai J, Uno N, Kazuki Y, Oshimura M, Minami I and Ikeda U. Allogeneic transplantation of iPS cell-derived cardiomyocytes regenerates primate hearts. *Nature*. 2016;538:388-391.

37. Soh BS, Ng SY, Wu H, Buac K, Park JH, Lian X, Xu J, Foo KS, Felldin U, He X, Nichane M, Yang H, Bu L, Li RA, Lim B and Chien KR. Endothelin-1 supports clonal derivation and expansion of cardiovascular progenitors derived from human embryonic stem cells. *Nat Commun*. 2016;7:10774.

38. Cai CL, Liang X, Shi Y, Chu PH, Pfaff SL, Chen J and Evans S. Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. *Dev Cell*. 2003;5:877-89.

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

40. Sun Y, Liang X, Najafi N, Cass M, Lin L, Cai CL, Chen J and Evans SM. Islet 1 is expressed in distinct cardiovascular lineages, including pacemaker and coronary vascular cells. *Dev Biol*. 2007;304:286-96.

41. Ahlgren U, Pfaff SL, Jessell TM, Edlund T and Edlund H. Independent requirement for ISL1 in formation of pancreatic mesenchyme and islet cells. *Nature*. 1997;385:257-60.

42. Pfaff SL, Mendelsohn M, Stewart CL, Edlund T and Jessell TM. Requirement for LIM homeobox gene Isl1 in motor neuron generation reveals a motor neuron-dependent step in interneuron differentiation. *Cell*. 1996;84:309-20.

43. Ghazizadeh Z, Fattahi F, Mirzaei M, Bayersaikhan D, Lee J, Chae S, Hwang D, Byun K, Tabar MS, Taleahmad S, Mirshahvaladi S, Shabani P, Fonoudi H, Haynes PA, Baharvand H, Aghdami N, Evans T, Lee B and Salekdeh GH. Prospective Isolation of ISL1. *Stem Cell Reports*. 2018;10:848-859.

44. Bruneau BG, Logan M, Davis N, Levi T, Tabin CJ, Seidman JG and Seidman CE. Chamber-specific cardiac expression of Tbx5 and heart defects in Holt-Oram syndrome. *Dev Biol*. 1999;211:100-8.

45. Chapman DL, Garvey N, Hancock S, Alexiou M, Agulnik SI, Gibson-Brown JJ, Cebra-Thomas J, Bollag RJ, Silver LM and Papaioannou VE. Expression of the T-box family genes, Tbx1-Tbx5, during early mouse development. *Dev Dyn*. 1996;206:379-90.

46. Liberatore CM, Searcy-Schrick RD and Yutzey KE. Ventricular expression of tbx5 inhibits normal heart chamber development. *Dev Biol*. 2000;223:169-80.

47. Sizarov A, Devalla HD, Anderson RH, Passier R, Christoffels VM and Moorman AF. Molecular analysis of patterning of conduction tissues in the developing human heart. *Circ Arrhythm Electrophysiol*. 2011;4:532-42.

48. Sizarov A, Anderson RH, Christoffels VM and Moorman AF. Three-dimensional and molecular analysis of the venous pole of the developing human heart. *Circulation*. 2010;122:798-807.

49. Später D, Abramczuk MK, Buac K, Zangi L, Stachel MW, Clarke J, Sahara M, Ludwig A and Chien KR. A HCN4+ cardiomyogenic progenitor derived from the first heart field and human pluripotent stem cells. *Nat Cell Biol*. 2013;15:1098-106.

50. Xie L, Hoffmann AD, Burnicka-Turek O, Friedland-Little JM, Zhang K and Moskowitz IP. Tbx5-hedgehog molecular networks are essential in the second heart field for atrial septation. *Dev Cell*. 2012;23:280-91.

51. Hiroi Y, Kudoh S, Monzen K, Ikeda Y, Yazaki Y, Nagai R and Komuro I. Tbx5 associates with Nkx2-5 and synergistically promotes cardiomyocyte differentiation. *Nat Genet*. 2001;28:276-80.

52. Lian X, Zhang J, Azarin SM, Zhu K, Hazeltine LB, Bao X, Hsiao C, Kamp TJ and Palecek SP. Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/β-catenin signaling under fully defined conditions. *Nat Protoc*. 2013;8:162-75.

53. Wu J, Anczuków O, Krainer AR, Zhang MQ and Zhang C. OLego: fast and sensitive mapping of spliced mRNA-Seq reads using small seeds. *Nucleic Acids Res*. 2013;41:5149-63.

54. Liao Y, Smyth GK and Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30:923-30.

55. Skelton RJ, Costa M, Anderson DJ, Bruveris F, Finnin BW, Koutsis K, Arasaratnam D, White AJ, Rafii A, Ng ES, Elefanty AG, Stanley EG, Pouton CW, Haynes JM, Ardehali R, Davis RP, Mummery CL and Elliott DA. SIRPA, VCAM1 and CD34 identify discrete lineages during early human cardiovascular development. *Stem Cell Res*. 2014;13:172-9.

56. Mommersteeg MT, Hoogaars WM, Prall OW, de Gier-de Vries C, Wiese C, Clout DE, Papaioannou VE, Brown NA, Harvey RP, Moorman AF and Christoffels VM. Molecular pathway for the localized formation of the sinoatrial node. *Circ Res*. 2007;100:354-62.

57. Protze SI, Liu J, Nussinovitch U, Ohana L, Backx PH, Gepstein L and Keller GM. Sinoatrial node cardiomyocytes derived from human pluripotent cells function as a biological pacemaker. *Nat Biotechnol*. 2017;35:56-68.

58. Birket MJ, Ribeiro MC, Verkerk AO, Ward D, Leitoguinho AR, den Hartogh SC, Orlova VV, Devalla HD, Schwach V, Bellin M, Passier R and Mummery CL. Expansion

and patterning of cardiovascular progenitors derived from human pluripotent stem cells. *Nat Biotechnol*. 2015;33:970-9.

59. Shiba Y, Filice D, Fernandes S, Minami E, Dupras SK, Biber BV, Trinh P, Hirota Y, Gold JD, Viswanathan M and Laflamme MA. Electrical Integration of Human Embryonic Stem Cell-Derived Cardiomyocytes in a Guinea Pig Chronic Infarct Model. *J Cardiovasc Pharmacol Ther*. 2014;19:368-381.

60. Ardehali R, Ali SR, Inlay MA, Abilez OJ, Chen MQ, Blauwkamp TA, Yazawa M, Gong Y, Nusse R, Drukker M and Weissman IL. Prospective isolation of human embryonic stem cell-derived cardiovascular progenitors that integrate into human fetal heart tissue. *Proc Natl Acad Sci U S A*. 2013;110:3405-10.

61. Shiba Y, Fernandes S, Zhu WZ, Filice D, Muskheli V, Kim J, Palpant NJ, Gantz J, Moyes KW, Reinecke H, Van Biber B, Dardas T, Mignone JL, Izawa A, Hanna R, Viswanathan M, Gold JD, Kotlikoff MI, Sarvazyan N, Kay MW, Murry CE and Laflamme MA. Human ES-cell-derived cardiomyocytes electrically couple and suppress arrhythmias in injured hearts. *Nature*. 2012;489:322-5.

62. Ye J, Gaur M, Zhang Y, Sievers RE, Woods BJ, Aurigui J, Bernstein HS and Yeghiazarians Y. Treatment with hESC-Derived Myocardial Precursors Improves Cardiac Function after a Myocardial Infarction. *PLoS One*. 2015;10:e0131123.

63. Fernandes S, Chong JJH, Paige SL, Iwata M, Torok-Storb B, Keller G, Reinecke H and Murry CE. Comparison of Human Embryonic Stem Cell-Derived Cardiomyocytes, Cardiovascular Progenitors, and Bone Marrow Mononuclear Cells for Cardiac Repair. *Stem Cell Reports*. 2015;5:753-762.

64. Moskowitz IP, Kim JB, Moore ML, Wolf CM, Peterson MA, Shendure J, Nobrega MA, Yokota Y, Berul C, Izumo S, Seidman JG and Seidman CE. A molecular pathway including Id2, Tbx5, and Nkx2-5 required for cardiac conduction system development. *Cell*. 2007;129:1365-76.

65. Harrild D and Henriquez C. A computer model of normal conduction in the human atria. *Circ Res*. 2000;87:E25-36.

66. Liu J, Laksman Z and Backx PH. The electrophysiological development of cardiomyocytes. *Adv Drug Deliv Rev*. 2016;96:253-73.

67. Uosaki H, Fukushima H, Takeuchi A, Matsuoka S, Nakatsuji N, Yamanaka S and Yamashita JK. Efficient and scalable purification of cardiomyocytes from human embryonic and induced pluripotent stem cells by VCAM1 surface expression. *PLoS One*. 2011;6:e23657.

68. Inagaki N, Hayashi T, Arimura T, Koga Y, Takahashi M, Shibata H, Teraoka K, Chikamori T, Yamashina A and Kimura A. Alpha B-crystallin mutation in dilated cardiomyopathy. *Biochem Biophys Res Commun*. 2006;342:379-86.

69. Rajasekaran NS, Firpo MA, Milash BA, Weiss RB and Benjamin IJ. Global expression profiling identifies a novel biosignature for protein aggregation R120GCryAB cardiomyopathy in mice. *Physiol Genomics*. 2008;35:165-72.

70. Kumarapeli AR, Su H, Huang W, Tang M, Zheng H, Horak KM, Li M and Wang X. Alpha B-crystallin suppresses pressure overload cardiac hypertrophy. *Circ Res*. 2008;103:1473-82.

71. Mitra A, Basak T, Datta K, Naskar S, Sengupta S and Sarkar S. Role of α-crystallin B as a regulatory switch in modulating cardiomyocyte apoptosis by mitochondria or endoplasmic reticulum during cardiac hypertrophy and myocardial infarction. *Cell Death Dis*. 2013;4:e582.

72. Zhou X, Chen JC, Liu Y, Yang H, Du K, Kong Y and Xu XH. Plasma Corin as a Predictor of Cardiovascular Events in Patients With Chronic Heart Failure. *JACC Heart Fail*. 2016;4:664-9.

73. Feistritzer HJ, Klug G, Reinstadler SJ, Reindl M, Mayr A, Mair J and Metzler B. Novel biomarkers predicting cardiac function after acute myocardial infarction. *Br Med Bull*. 2016;119:63-74.

74. Dong N, Chen S, Yang J, He L, Liu P, Zheng D, Li L, Zhou Y, Ruan C, Plow E and Wu Q. Plasma soluble corin in patients with heart failure. *Circ Heart Fail*. 2010;3:207-11.

75. Yan W, Wu F, Morser J and Wu Q. Corin, a transmembrane cardiac serine protease, acts as a pro-atrial natriuretic peptide-converting enzyme. *Proc Natl Acad Sci U S A*. 2000;97:8525-9.

76. Wu Q. The serine protease corin in cardiovascular biology and disease. *Front Biosci*. 2007;12:4179-90.

77. Zhou Y and Wu Q. Corin in natriuretic peptide processing and hypertension. *Curr Hypertens Rep*. 2014;16:415.

78. Constam DB and Robertson EJ. SPC4/PACE4 regulates a TGFbeta signaling network during axis formation. *Genes Dev*. 2000;14:1146-55.

Chapter 2:

Direct Cardiac Reprogramming

Abstract

The human adult heart lacks a robust endogenous repair mechanism to fully restore cardiac function after insult, thus the ability to regenerate and repair the injured myocardium remains a top priority in treating heart failure. The ability to efficiently generate a large number of functioning cardiomyocytes capable of functional integration within the injured heart has been difficult. However, the ability to directly convert fibroblasts into cardiomyocyte-like cells both *in vitro* and *in vivo* offers great promise in overcoming this problem.

Introduction

Heart failure (HF), the leading cause of death and hospitalizations worldwide, results from a myriad of cardiovascular diseases that lead to the death or dysfunction of cardiomyocytes. With a prevalence of 38 million people worldwide, it places a significant financial burden on health care systems, with an estimated $30 billion of annual spending in just the United States alone [1, 2]. Despite recent advances in the care and management of HF patients, the prognosis of advanced HF remains dismal at 50% survival at 5 years, a rate often lower than many cancers [3, 4]. Considering that the pathophysiology of HF involves death or dysfunction of the cardiac myocyte, new therapeutic strategies for heart regeneration may offer hope to this intractable disease.

The human adult heart lacks endogenous repair mechanisms to fully restore cardiac function after an insult, thus the ability to regenerate and repair the injured myocardium remains a top priority in treating HF. However, the ability to efficiently generate a large number of functioning cardiomyocytes capable of functional integration within the injured heart has remained an obstacle. Current cell therapies are focused on three main approaches; 1) induction of endogenous cardiomyocytes to undergo proliferation and repopulate the damaged myocardium, 2) transplantation of cardiovascular progenitor cells (CPCs) or cardiomyocytes generated through the differentiation of pluripotent stem cells, and 3) direct reprogramming of somatic cells to cardiomyocytes or expandable CPCs without transitioning through a pluripotent intermediate. This chapter is focused on the last approach.

Direct reprogramming was first reported in 1987 when a single cDNA encoding

MyoD was transfected into fibroblasts converting them into muscle myoblasts [5]. A few years later MyoD was identified as the master regulator gene for skeletal muscle development [6]. The ability to directly reprogram adult cells to a desirable fate demonstrate an immense potential of this powerful tool for tissue regeneration and replacement. Since the identification of MyoD, there has been extensive focus on identification of master regulator(s) for other cell lineages and this search has led to the successful conversion of mature cells into other cells types including myoblasts, neurons, hepatocytes, intestinal cells, blood progenitor cells, and cardiomyocytes [5, 7-11].

In this chapter, the insights and progress that has been gained from investigation of direct cardiac reprogramming, with a focus on the use of key transcription factors and other cardiogenic genes are reviewed. Furthermore, the use of other biologics and small molecules to improve the efficiency of cardiac reprogramming and the development of safe reprogramming approaches for clinical application are explored.

Reprogramming of Somatic Cells to Cardiomyocyte-like Cells by Over-expression of Key Cardiac Transcription Factors

Direct reprogramming of fibroblasts into cardiomyocyte-like cells was first reported in 2010 using viral overexpression of three important cardiac developmental transcription factors (TFs), *Gata4*, *Mef2c*, and *Tbx5* (GMT) in mouse cardiac and tail-tip fibroblasts [11]. The Srivastava group used an iterative screening approach in which 14 factors were removed one by one to identify those that were dispensable for direct reprogramming. This process ultimately identified GMT as the factors sufficient to induce conversion of fibroblasts to cardiomyocyte-like cells without transitioning through a progenitor state.

TBX5 is an important T-box TF involved in early cardiac development that directs formation of the primary heart field through a coordinated but yet complex interaction with other TFs [12]. One such interaction is with GATA4, a member of the GATA family zinc-finger TFs, which modifies the chromatin structure allowing other TFs such as NKX2-5 to bind their targets and fully activate the cardiac transcriptional program [13]. MEF2c, a MADS box transcription enhancer factor, is important for the formation of the secondary heart field through its interaction with other cardiac TFs [14]. After the establishment of GMT as the core TFs for direct cardiac reprograming, much of the focus transitioned to improving the reprogramming efficiency and/or the function of the induced cardiomyocyte-like cells (iCMs) through addition of other important cardiac TFs to GMT. This was mainly due to the poor efficiency of reprogramming, reported to be 4.8% cardiac troponin T+ (cTnT+) cells in the original paper. Additionally, it was soon noted that GMT alone was insufficient to convert human fibroblasts to iCMs.

One of the first TFs added to GMT was the bHLH TF HAND2 (referred to as GMHT). In cardiac development, HAND2 plays an important role in the formation of the ventricular chambers through interaction with GATA4 and NKX2-5 [15]. GMHT treatment of mouse embryonic fibroblasts (MEFs) resulted in iCMs expressing low levels of sarcomeric proteins and displayed immature characteristics of the main cardiac cell types (atrial, ventricular, and pacemaker) [16]. In an effort to increase transcriptional activity of GMHT, the transactivation domain of MyoD was fused to each G, M, H or T and over-expressed in mouse fibroblasts. When *MyoD* was fused to *Mef2c*, a 15-fold increase in reprogramming efficiency was observed [17]. Other TFs that are essential during cardiovascular development have also been studied for direct reprogramming. NKX2-5,

a homeobox TF critical for normal heart morphogenesis, was overexpressed in mouse fibroblasts in addition to GMHT. This combination resulted in more than 50-fold increase in the efficiency of cardiac reprogramming compared to GMT alone and produced iCMs with mature cardiomyocyte marker expression, robust calcium oscillation, and spontaneous beating [18].

Additionally, an alternative screening approach that surveyed triplet combinations of 10 important cardiac TFs revealed that *Tbx5*, *Mef2c* and *Myocd*, a developmental regulator of cardiomyocytes and smooth muscle cells, were able to induce a more complete cardiac phenotype than GMT in mouse fibroblasts [19]. Likewise, a combinatorial screen of 10 transcriptions factors added to GMT in MEFs identified a combination of cocktails that resulted in successful reprogramming. GMT plus *Myocd* and *Srf*, a TF important in mesoderm formation, or GMT plus *Myocd*, *Srf*, *Mesp1*, another mesoderm inducing TF and *Smarcd3*, a chromatin structure altering protein, enhanced reprogramming and the expression of cardiac sarcomeric proteins over GMT alone [20].

Despite the successes of TF overexpression to reprogram murine cells, similar approaches to reprogram human somatic cells has been more difficult to achieve. Only a few studies have reported successful reprogramming of human cells to iCMs using TFs alone. The first of these studies reported a combination of the E26 transformation-specific (ETS) TF family member *ETS2* and *MESP1* proteins to induce reprogramming of human dermal fibroblasts to cardiac progenitors [21]. Another study using GMT with *MESP1* and *MYOCD* in human cardiac and dermal fibroblasts was sufficient to induce expression of multiple cardiac-specific proteins, increased a broad range of cardiac genes, and exhibited spontaneous calcium transients [22]. A third report showed that expressing GMT

along with *ESSRG* (a transcriptional activator), *MESP1*, *MYOCD*, and *ZFPM2* (a modulator of GATA proteins) in human fetal cardiac fibroblasts and neonatal skin fibroblasts enhanced cardiac reprogramming, sarcomere formation, calcium transients, and action potentials [23]. Results of TF based reprogramming are summarized in the table below.

Table 4: Summary of transcription factor direct cardiac reprogramming results

Reprogramming Factors	Mouse/ Human	*in vitro/ in vivo*	Reported Efficiency	Analysis Method	Reference
GMT	Mouse	Both	4.8% cTnT+ *(in vitro*) 17% αMHC+ *(in vitro*)	FC	11, 57
GMHT	Mouse	Both	27.6% cTnT+ *(in vitro*)	FC	16, 53
GMHT, *MyoD* transactivation domain	Mouse	*in vitro*	19% cTnT+	IF	17
GMHT, *Nkx2-5*	Mouse	*in vitro*	1.6% GCaMP+	IF	18
MT, *Myocd*	Mouse	*in vitro*	12% cTnT+	FC	19
GMT, *Myocd*, *Srf*, *Mesp1*, *Smarcd3*	Mouse	*in vitro*	2.4% αMHC+	FC	20
ETS2*, *MESP1	Human	*in vitro*	13.7% αMHC+	FC	21
GMT, *MESP1*, *MYOCD*	Human	*in vitro*	5.9% cTnT+	FC	22
GMT, *ESSRG*, *MESP1*, *MYOCD*, *ZFPM2*	Human	*in vitro*	18.1% αMHC+	FC	23

Abbreviations: FC: Flow cytometry IF: Immunofluorescence

Improving the Efficiency of Direct Reprogramming with Biological Molecules

Despite the successes of direct reprogramming using forced expression of cardiac TFs, the efficiency remains low. The initial report on direct conversion of fibroblasts to

cardiomyocyte-like cells noted an efficiency of 4.8%. In an effort to improve reprogramming efficiency, many methods have been developed using additional molecules. These additives can be classified into three major categories: inhibitors/cytokines, non-coding RNAs, and epigenetic modifiers. A summary of these reprogramming experiments is presented in Table 5.

Inhibitors/cytokines

A potential approach to improving reprogramming is to inhibit the endogenous signaling pathways and gene programs that maintain the distinct properties of fibroblasts. One of the major signaling pathways active in fibroblasts is the transforming growth factor (TGF)-β pathway. TGF-β has diverse and pleotropic effects through its activation and signaling. The downstream effect of TGF-β signaling pathway involves phosphorylation of receptor-regulated SMADs that ultimately activate TFs that participate in regulation of target gene expressions, many of which are critical in fibroblast activation and proliferation. Since inhibition of TGF-β has been shown to increase mouse embryonic stem cell differentiation to cardiomyocytes, [24, 25] it was hypothesized that TGF-β inhibition could improve reprogramming. The TGF-β inhibitors SB431542 and A83-01 have been added to various reprogramming combinations and have shown an increase in reprogramming efficiency. SB431542 is a selective and potent inhibitor of the TGF-β pathway through suppression of the Activin A receptors ALK5, ALK4 and ALK7. A83-01 is also a selective inhibitor of ALK5, ALK4 and ALK7, but is more potent than SB431542 in its inhibition and effectively blocks phosphorylation of Smad2. When SB431542 was combined with GMHT, a 5-fold increase in reprogramming efficiency was observed in both MEFs and mouse adult cardiac fibroblasts [26]. Furthermore, when GMT and

SB431542 were combined with WNT inhibition by XAV939, reprogramming efficiency was increased 8-fold in cardiac fibroblasts compared to GMT alone [27]. In addition to the TGF-β pathway, other pro-fibrotic and intracellular signaling pathways such as the Rho-associated kinase, JAK/STAT, Notch and Akt pathways have been targeted to improve reprogramming [28-31].

The utility of other molecules to enhance cardiac reprogramming have been inspired by using cytokines and/or modulators that are critical in mammalian cardiac development, many of which are commonly used in the differentiation of cardiomyocytes from pluripotent stem cells. Fibroblast growth factor-2 (FGF2), FGF10, and vascular endothelial growth factor (VEGF) in combination with GMT or GMHT in MEFs and mouse tail-tip fibroblasts showed an increase in the number of iCMs that spontaneously contract [32]. These approaches also accelerated maturation of iCMs *in vitro* and thus, the activation of important developmental pathway during reprogramming warrants further research.

Non-coding RNAs

MicroRNAs (miRNAs) are small non-coding RNAs that induce degradation or inhibit translation of target mRNAs. miRNAs are an attractive additive to reprogramming since they play important roles in the post-transcriptional regulation of cardiac gene expression and have critical function in almost every stage of heart development. miRNA-1, miRNA-133, miRNA-208, and miRNA-499 have been shown to be cardiac- and muscle- specific and play important roles in cardiac development and function. miRNA-1, contributing to ~40% of total miRNAs in the heart, has been shown to promote cardiomyocyte proliferation and suppress apoptosis during development, however its

function in cardiac reprogramming remains unknown [33]. miRNA-133 is important in orchestrating cardiac development, gene expression, growth, and function [34]. It also promotes cardiomyocyte proliferation by repressing the transcriptional regulator SNAI1 and silences fibroblast gene signatures during reprogramming [35]. These two miRNAs have been used in combination with other factors to successfully enhance cardiac reprogramming. A novel approach was reported recently where a combination of miRNAs promoted direct conversion of cardiac fibroblasts into cardiomyocyte like cells without the need for forced expression of exogenous TFs. A combination of miRNAs-1, -133, -208, and -499 was reported to be sufficient to convert mouse cardiac fibroblasts into iCMs without addition of other factors *in vivo* [36]. The potential mechanism for this effect is thought to be due to altered H3K27 methyltransferase and demethylase expression, which leads to changes in the epigenetic landscape of fibroblasts to induce their conversion into cardiomyocyte-like cells. A 10-fold increase in miRNA-mediated murine cardiac fibroblast reprogramming was observed when miRNAs -1, -133, -208, and -499 were combined with JAK Inhibitor I [30].

Similarly, when miRNA-133 was used in conjunction with GMT, *Mesp1*, and *Myocd* or GHT, *Myocd*, and miRNA-1, the reprogramming efficiency was increased in both human and mouse fibroblasts by repressing Snai1 and silencing fibroblast gene signatures [35, 37]. Zhao et al used a combination of GMHT, miRNA-1, -133, -208, -499, Y-27632 and A83-01 in MEFs and mouse adult fibroblasts to achieve ~60% Cardiac Troponin T+ and 60% α-actinin+ iCMs [29]. miRNA-590, a miRNA that can induce adult cardiomyocyte proliferation, was recently shown to be able to replace *HAND2* and *MYOCD* in GMT direct reprogramming experiments using human and porcine fibroblasts

[38, 39]. While GMT was initially shown to be sufficient for cardiac reprogramming, further studies have indicated that a multi-prong approach may be necessary to enhance reprogramming and could hold great promise for future *in vivo* clinical application.

Long non-coding RNAs (lncRNAs) are a heterogeneous group of transcripts longer than 200 nucleotides that exert major regulatory roles in gene expression during development and disease through many different mechanisms. Recent advances in sequencing and analysis technologies have allowed many lncRNAs to be identified, but due to their complex and multiple mechanisms of action as well as the low interspecies conservation it has been difficult to decipher biological functions of many lncRNAs [40]. A list of cardiac lncRNAs that are involved in cardiac differentiation, development and contractile function have been reviewed [41]. Braveheart (*Bvht*) and *Fendrr* play a critical role in cardiac lineage commitment by regulating the transition from mesoderm to CPCs through activation of key cardiac development genes and TFs including some of those studied above [42-44]. *Hotair*, *Chaer* and other lncRNAs have also been shown to regulate the epigenetic landscape in cardiac development by regulating proteins involved in histone modification at targeted promoters [45, 46]. LncRNAs play an extensive role in the regulation of cardiac development and gene expression therefore it may be advantageous to explore the use of lncRNAs in direct reprogramming studies, however no direct reprogramming studies have been published using lncRNAs.

Epigenetic modifiers

Reprogramming of one somatic cell type to another requires the activation and repression of multiple sets of genes, leading to vast genomic changes. The epigenetic

landscape plays an important role in determining the reprogramming efficiency as accessibility of TFs to their DNA targets is critical. During reprogramming, epigenetic marks such as histone methylation, acetylation and ubiquitination must be added and removed from fibroblast- and cardiac- specific genes. These modifications will suppress expression of fibroblast genes while activating cardiac genes by remodeling chromatin structure to allow or restrict access of TFs to their target genes. It has been shown that during cardiac direct reprogramming the trimethylated histone H3 of lysine 27 (H3K27me3), which marks inactive chromatin increases at fibroblast promoters and decreases at cardiac promoters while the activated chromatin mark H3K4me3 shows the opposite pattern at important loci [11, 47]. Moreover, the activating H3K4me2 histone mark has been shown to be increased at the regulatory regions of miRNA-1 and -133 [29]. To this end, attempts to improve direct cardiac reprogramming have been carried out using modulators of epigenetic marks. Bmi1 was identified as a barrier to reprogramming by modifying histone marks at key cardiogenic loci, thus inhibiting iCM induction. When *Bmi1* activity was knocked down, the active histone mark, H3K4me3, was increased while the repressive H2AK119ub mark was reduced, leading to increased cardiac gene expression at important loci [48]. In non-integrative and *in vivo* reprogramming experiments discussed later in this review, other epigenetic modifiers that inhibit histone methyl transferases and histone demethylase have been used. The importance of epigenetic landscape and changes that happen during reprogramming have recently begun to be unraveled using a single-cell transcriptomic approach by the Qian group [49]. These results highlight the complexity of the reprogramming process and the importance influence of a variety of factors, warranting additional research into the sequential addition of TFs, non-coding

RNAs, cytokines, inhibitors, and epigenetic modifiers to further improve the reprogramming efficiency.

Modified RNAs

Another promising non-viral method of direct cardiac reprogramming is the use of modified mRNAs (modRNAs) [50]. ModRNAs are non-cytopathic and do not integrate into the host genome, thus offering a safer approach to reprogramming. ModRNAs have been used successfully to generate induced pluripotent stem cells from somatic cells through transient expression of mRNAs that direct cell fate. ModRNAs are produced using an *in vitro* transcription system to generate mRNAs that contain a synthetic 5' guanine cap and poly-A tail, which improves half-life and stability, as well as modified nucleotide bases that reduce the innate immune response of the host cell. This technology is endowed with a number of attractive properties that would make it a potentially powerful platform for direct cardiac reprogramming. ModRNAs can mediate robust and dose-titratable expression of key TFs over a specified time and in a particular sequence. Previous studies outlined in this review have highlighted the fact that direct cardiac reprogramming is a complex process that may require sequential treatments to better overcome the reprogramming barrier. ModRNAs may be ideal for direct reprogramming as they have a relatively short half-life therefore distinct factors can be expressed for a short period of time then removed from the reprogramming cocktail or added again to continue expression. ModRNAs may open the door to following a more developmentally relevant sequence of TFs to improve transcription. It is also foreseeable that modRNAs could be combined with other small molecules, cytokines, and non-coding RNAs discussed in this chapter.

Table 5: Summary of direct cardiac reprogramming results

Reprogramming Factors	Mouse/ Human	*in vitro/ in vivo*	Reported Efficiency	Analysis Method	Reference
GMHT, SB431542	Mouse	*in vitro*	9.3% GCaMP+	IF	26
GMHT, DAPT	Mouse	*In vitro*	38% cTnT+	IF	28
GM(H)T, FGF2, FGF10, VEGF	Mouse	*in vitro*	2.9% cTnT+	FC	32
miRNA-1, -133, -208, -499	Mouse	*in vivo*	12% tdTomato+cTnT+	IHC	36
miRNA-1, -133, -208, -499, JAK Inhibitor I	Mouse	Both	28% αMHC+ *(in vitro)*	FC	30
GMT, *Mesp1*, *Myocd*, miRNA-133	Mouse	*in vitro*	12.9% cTnT+	FC	35
	Human		27.8% cTnT+		
GHT, *MYOCD*, miRNA-1, -133	Human	*in vitro*	34.1% cTnT+	FC	37
GMHT, miRNA-1, -133, -208, -499, Y-27632, A83-01	Mouse	*in vitro*	60% cTnT+	IF	29
GMT, miRNA-590	Human	*in vitro*	4.6% cTnT+	FC	39
	Porcine				
Ascorbic acid, RepSox, forskolin, valproic acid, CHIR99021	Mouse	*in vitro*	9% αMHC+	FC	62
CHIR99021, BIX01294, A83-01, AS8351, SC1, OAC2, Y27632, SU16F and JNJ10198409	Human	*in vitro*	6.6% cTnT+	FC	63

Abbreviations: FC: Flow cytometry IF: Immunofluorescence IHC: Immunohistochemistry

Methods

Production of modRNAs

ModRNAs were generated using the methods outlined by Mandal and Rossi [51]. DNA sequences for genes of interest were cloned into the MCS of the pMRNA vector (SBI), linearized by restriction enzyme digestion and purified by gel elctrophoresis. Next, a tail PCR was performed to add a poly-A tail to the cDNA template using a poly dT primer. The resulting PCR product was purified by gel electrophoresis and then used in an *in vitro* transcription reaction using a RNA cap analog (3′-O-Me-m^{7}G(5′)ppp(5′)G, Trilink), modified nucleotides (5-Methylcytidine-5′-triphosphate (Me-CTP, Trilink) and Pseudouridine-5′-triphosphate (Pseudo-UTP; Trilink)) and recombinant T7 RNA polymerase (MEGAscript T7 kit, Ambion) to produce the modified RNA. The modified RNA was then treated with DNase (NEB) and Antarctic phosphatase (NEB) then spin column purified (MEGAclear kit, Ambion) and quantified using Nanodrop.

Isolation of mouse embryonic fibroblasts

MEFs were isolated from αMHC-GFP+ pups as previously described [52]. Briefly E13.5 embryos were isolated from pregnant dams and analyzed for GFP expression in the heart by fluorescence microscopy. The heads and red tissue (heart and lungs) were surgically removed from GFP+ embryos then the remaining tissue was chopped into 1-2 mm pieces using a razor blade. The embryo pieces were then transferred to a tube containing 0.25% Trypsin-EDTA (Thermo Fisher) and incubated for 10 minutes at 37°C. The pieces were mixed by pipetting and incubated for an additional 5-10 minutes. 20 ml of MEF culture media (DMEM containing GlutaMAX, 10% FBS and 1% Pen/Strep) was

added and large pieces were allowed to settle by gravity for 5 minutes before the remaining supernatant was transferred to a T75 flask (3 embryos per flask) coated with 0.1% gelatin. Upon reaching confluency the MEFs were verified to be GFP- by microscopy and frozen for subsequent experiments.

Direct cardiac reprograming

To generate iCMs, fibroblasts were transduced with GMT or GMHT retroviruses as previously described [53]. In brief, isolated fibroblasts from the first passage were plated on gelatin (0.1%) coated dishes at 7500 cells/cm^2. 24 hours later media was changed and 5 µM SB431542 or 0.5 µM A83-01 (Cayman Chemical) was added and kept throughout reprogramming. The next day fibroblasts were transduced for 24 hours with retroviruses containing 6 µg/ml polybrene (Sigma) or 500 ng of each modRNA was transfected using Lipofectamine MessengerMAX reagent (Thermo Fisher) according to the manufactures protocol. Following transduction of transfection the fibroblasts were then transferred to iCM media (10% FBS, 5% horse serum, 1% penicillin/streptomycin, 0.1 mm non-essential amino acids (Life Technologies), essential amino acids (Life Technologies), B-27 supplement (Life Technologies), insulin-selenium-transferrin (Life Technologies) and vitamin mixture (Sigma) in DMEM/M199 (4:1)). Media was changed every 2-3 days.

Animal Use and Care

Animal housing, handling, and procedures were approved by and carried out in accordance with guidelines set by the UCLA Animal Research Committee (ARC), UCLA Institutional Animal Care and Use Committee (IACUC), and the NIH Guide for the Care and Use of Laboratory Animals.

Results

Generation and validation of ModRNAs

ModRNAs are generated using an *in vitro* transcription system to generate mRNAs containing a synthethic 5' guanine cap and poly-A tail to increase half-life and stability as well as modified nucleotide bases to reduce the innate immune response of the host cell (Figure 16A). In order to test the feasibility of this method we generated a nuclear localized destabilized GFP modified RNA. Subsequently, 500ng of purified nuclear localized destabilized GFP modRNA was transfected into mouse embryonic fibroblasts using Lipofectamine MessengerMAX and analyzed 24 hours later for transfection efficiency based on GFP expression. We were able to successfully generate and transfect modRNA as indicated by GFP expression in the majority of mouse embryonic fibroblasts (Figure 16B). The transfection efficiency remains consistently at 75-85%, which suggests that this method can be used for reprogramming in order to maintain high expression levels. Additionally, modRNAs for the cardiac transcription factors, *Gata4*, *Mef2c* and *Tbx5* were generated and tested for expression of the modRNAs at the protein level 24 hours post transfection in HEK293T cells using immunocytochemistry for the specific cardiac transcription factors (Figure 16C). Our preliminary studies also have demonstrated that the expression levels of the modRNAs significantly decrease after 40 hours, making this an attractive feature for temporal expression of certain transcription factors for efficient reprogramming.

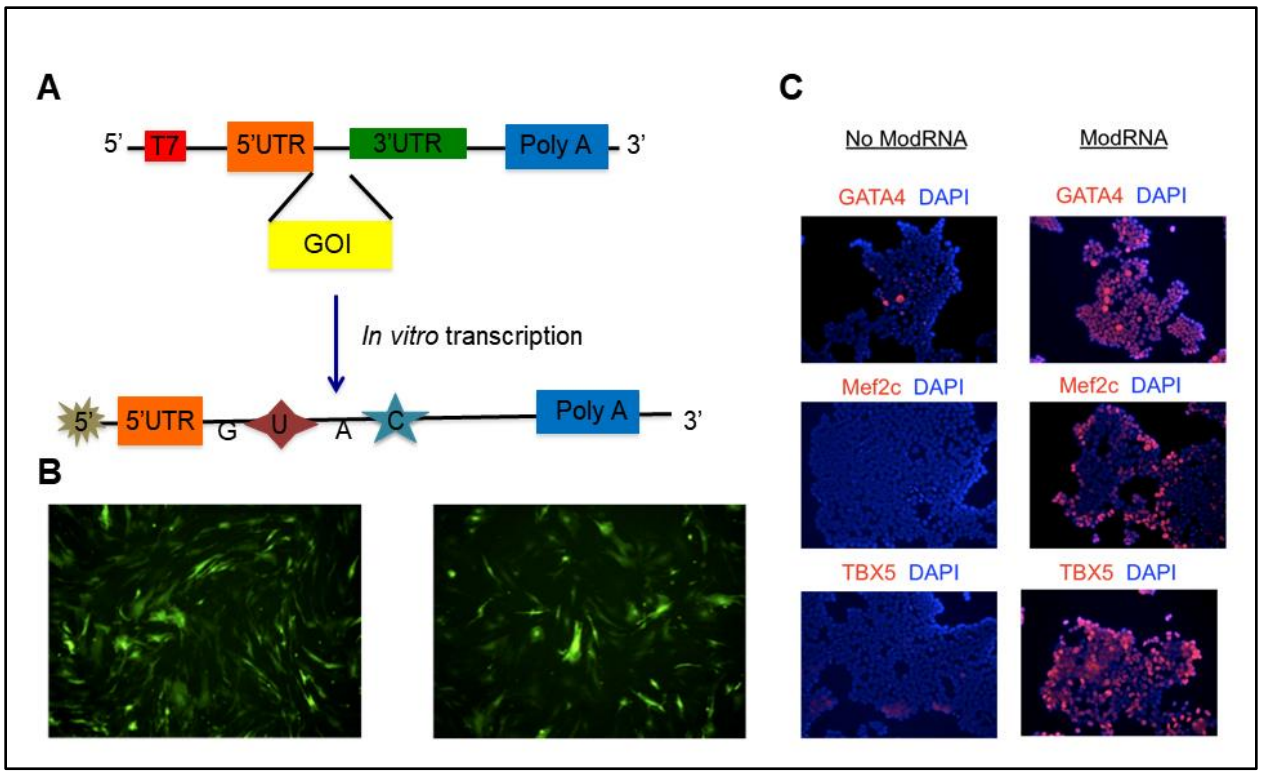


Figure 16. Generation and validation of modRNAs. (A) Schematic showing the *in vitro* production of modRNAs containing a 5' cap and modified Uracil and Cytosine bases (GOI: Gene of interest). (B) MEFs showing GFP expression 18 hours post transfection with GFP modRNA. (C) HEK293T immunocytochemistry showing expression of cardiac TFs 18 hours post transfection with GATA4, Mef2c and TBX5 modRNAs.

Direct reprogramming of mouse embryonic fibroblasts

To test and compare direct cardiac reprogramming efficiencies we isolated mouse embryonic fibroblasts (MEFs) from αMHC-GFP transgenic mice. These MEFs are colorless after isolation, but upon expression of the cardiac contractile protein, αMHC, after successful cardiac reprogramming will fluoresce green allowing for easily quantification by flow cytometry. Isolated αMHC-GFP MEFs were then either transduced with retroviruses expressing, GMT or GMHT, or transfected with GMT modRNAs and subjected to flow cytometry 7 days later. Flow cytometry analysis for αMHC-GFP and cardiac troponin T (cTnT), another more mature cardiac contraction protein, revealed that

modRNA cardiac reprogramming is possible albeit at a lower efficiency than retrovirus reprogramming (Figure 17A).

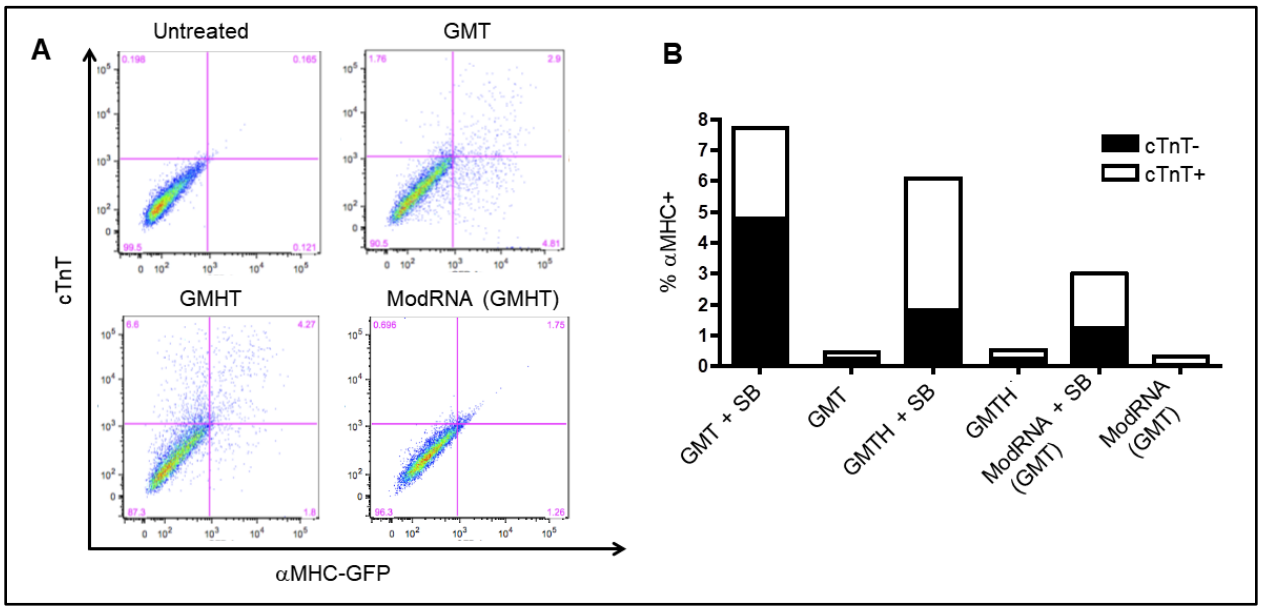


Figure 17. Direct reprogramming using modRNAs and enhancement with SB431542. (A) Day 7 flow cytometry analysis comparing retroviral and modRNA reprogramming. (B) Comparison of reprogramming efficiencies in the presence or absence of SB431542.

Next, we sought to improve the reprogramming efficiency by including the TGF-β inhibitor, SB431542, into our reprogramming conditions. Flow analysis at day 7 of reprogramming showed that the addition of SB431542 markedly improved the reprogramming efficiencies in all conditions including those using modRNAs (Figure 17B). Likewise, when we added the TGF-β inhibitor, A83-01, we observed an increase in reprogramming efficiency with frequent areas of spontaneous contraction and enhanced expression of cardiac contractile proteins (Figure 18A and 18B).

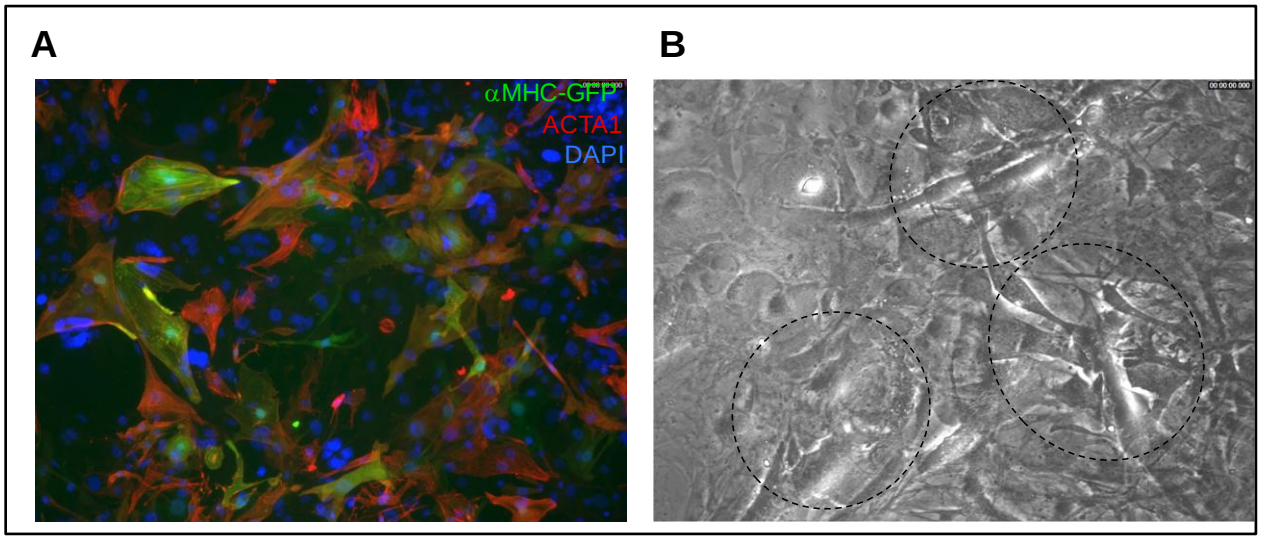


Figure 18. GMT + A83-01 reprogramming in MEFs. (A) Immunocytochemistry for the cardiac markers αMHC-GFP and a-actinin (ACTA1). (B) Video snapshot showing beating areas of reprogrammed cells (outlined areas).

Temporal modRNA direct reprogramming

One advantage of modRNAs for reprogramming is the relatively short half-life of these mRNAs which allow for temporal control of the expression of important TFs. To examine if temporal expression of cardiac TFs that more closely mimics development can improve the efficiency of direct reprogramming, we produced *Mesp1* modRNA. We hypothesized that by first priming the cells with an early mesoderm TF such as *Mesp1* then adding later cardiac TFs we could improve the reprogramming efficiency. Indeed, when we added *Mesp1* modRNA for 3 days followed by retroviruses for GMT or GMHT for 7 days then analyzed αMHC-GFP and cTnT expression by flow cytometry we observed an increase in reprogramming efficiency (Figure 19A). These data provide evidence that temporal control of reprogramming may play an important role in increasing reprogramming efficiencies. Furthermore, the ability of modRNAs to temporally control

expression and/or dosage of TFs is an attractive avenue to reprogramming that warrants further investigation.

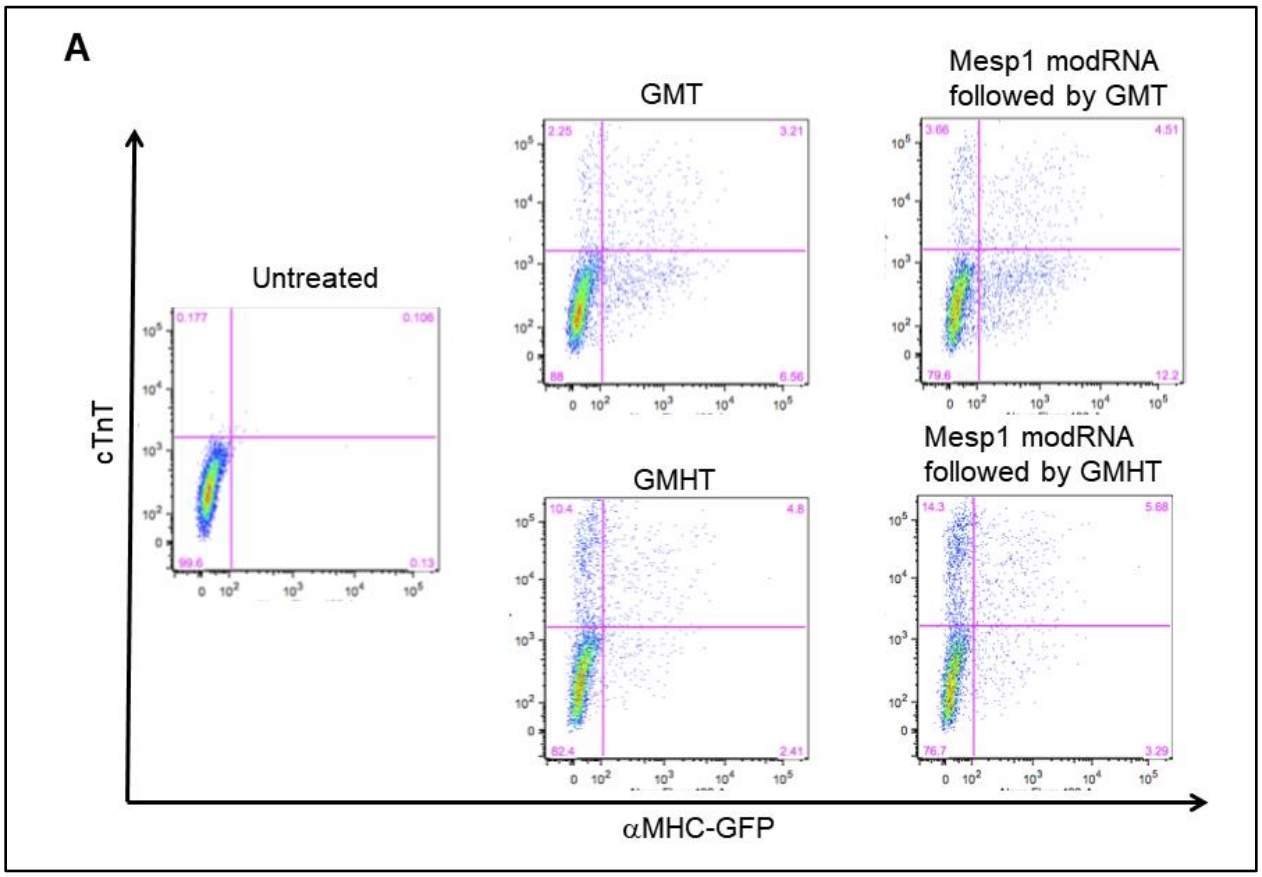


Figure 19. Temporal control of modRNAs enhance reprogramming efficiency. (A) Flow cytometry analysis at day 10 of reprogramming after 3 days treatment with *Mesp1* modRNA followed by GMT retroviruses.

Discussion

Direct Reprogramming to expandable cardiac progenitor cells

Another recent application of direct reprogramming is the generation of expandable CPCs. The goal of this approach is to safely create CPCs *in vitro* that can then be expanded in culture before transplantation into the injured heart. Upon transplantation, the CPCs will differentiate into three major cells of the heart; cardiomyocytes, endothelial and smooth muscle cells. Two groups were able to successfully generate expandable CPCs using unique reprogramming cocktails containing a variety of biomolecules described above. The Kamp group generated CPCs from fibroblasts using Mesp1, Tbx5, Gata4, Nkx2-5 and Baf60c, a chromatin remodeling protein. *In vitro* expansion and maintenance of a CPC state was achieved using a Wnt activator, BIO, and a JAK/STAT activator, LIF [54, 55]. On the other hand, the Ding group used a chemical approach to reprogramming fibroblasts to CPCs. Generation, expansion and maintenance of CPCs was achieved by addition of BMP4, Activin A, CHIR99021 and SU504, a FGF, VEGF and PDGF signaling inhibitor [56]. Both groups were able to show that their reprogrammed CPCs maintained their characteristics for many passages in culture and could generate cardiomyocytes, endothelial and smooth muscle cells both *in vitro* and *in vivo* when transplanted. Direct reprogramming of fibroblasts to CPCs represents a scalable method for the generation of multiple cardiac cell types for clinical applications, however this approach has not yet been applied to human cells.

Progress of *in vivo* Direct Reprogramming

The ultimate goal of direct reprogramming is to be able to repair the damaged myocardium after injury. Direct reprogramming offers two potential approaches for heart regeneration; 1) transplantation of reprogrammed fibroblasts (either iCMs or CPCs) into the infarcted heart and 2) reprogramming resident cardiac fibroblasts directly to cardiomyocytes. The first attempt at cardiac regeneration using direct reprogramming was carried out using cardiac fibroblasts that were transduced with GMT for 1 day and then transplanted into mouse hearts [11]. Analysis of these cells post-transplantation revealed that they successfully generated cardiomyocyte-like cells *in vivo*. Other studies have used *in vivo* transplantation of reprogrammed cells to test their regenerative potential. However, cell transplantation is complicated by many factors such as cell retention, viability, structural and functional integration, and immune rejection. Therefore, *in situ* repair of the heart is best studied by targeting endogenous cardiac fibroblasts through viral transfection of the infarct zone. This approach was attempted in 2012 by the Olson and Srivastava groups, in which local delivery of GM(H)T viruses induced reprogramming of non-myocytes into iCMs by 4 weeks post-surgery [53, 57]. Additionally, the Srivastava study reported that co-delivery of thymosin β4 and GMT viruses further improved ejection fraction and reduced scar formation. Interestingly, it has been reported that the *in vivo* cardiac niche may improve the efficiency of reprogramming, however, the mechanisms underlying this observation remain elusive [27, 36, 53, 57]. Several studies have improved the *in vivo* reprogramming efficiency even further by optimizing polycistronic expression vectors to control the stoichiometry of TF expression or by the addition of small molecules delivered with TFs [27, 58-60].

Non-integrative Methods of Direct Reprogramming for Future *In vivo* Applications

The reprogramming results shown thus far suggest that direct reprogramming of fibroblasts can be a feasible therapeutic approach to repairing the injured myocardium. However, relatively safe methods for the delivery of various reprogramming factors needs to be explored for *in vivo* applications. Adeno-associated virus (AAV) vectors are attractive tools for TF delivery, but the limited capacity of about 4.5 kb complicates the expression of multiple TFs in a single vector and still involves the use of an integrative viral system. Sendai virus reprogramming is an appealing alternative to AAV since it does not integrate into the host genome and has been successfully used to reprogram many different cell types to pluripotency, however, its use in direct reprogramming has not yet been explored. A recent study showed that expression of GMT in nonintegrating, acute expression adenoviral vectors were as efficient at reprogramming as lentiviral vectors in a rat infarct model, which has increased the clinical applicability of *in vivo* reprogramming. As described in two recent reports, the temporal and stoichiometric control of TFs are also important in determining reprogramming efficiency [60, 61]. Unfortunately, current *in vivo* viral reprogramming tools are unable to control dosage and temporal expression of TFs but warrant further investigation to improve reprogramming efficiency.

Compared to TFs and miRNAs, small molecules have many advantages such as more effective cell delivery, non-immunogenic, less expensive, and are generally safer. Moreover, it is more convenient to control the process of reprogramming through varying small molecule concentrations and combinations *in vitro*. A combination of ascorbic acid, RepSox (a TGF-β inhibitor), forskolin, valproic acid and CHIR99021 (a WNT pathway activator through the inhibition of glycogen synthase kinase 3) was shown to reprogram MEFs and mouse tail-tip fibroblasts to iCMs *in vitro* [62]. Ding and colleagues were able to

use a cocktail of 9 small molecules (CHIR99021, A83-01, BIX01294, AS8351 SC1, OAC2, Y27632, SU16F, and JNJ10198409) to direct cardiac reprogram of human foreskin fibroblasts *in vitro* before transplantation in injured murine hearts [63]. Among these small molecules were the epigenetic modifiers BIX01294 (a methyltransferase inhibitor) and AS8351 (a histone demethylase inhibitor), SC1 (an ERK2 and Ras-GAP inhibitor), OAC2 (an Oct4 activator), SU16F (a PDGFRβ inhibitor) and JNJ10198409 (a PDGF receptor tyrosine kinase inhibitor). However, the use of small molecules for *in vivo* reprogramming poses some unanswered questions. Small molecules can enter the blood stream and spread to other organs with unknown consequences. Additionally, the ability of timely uptake into specific target cell type remains a challenge. Development of novel biomaterials for local delivery, controlled release, and retention of small molecules are still needed.

Roadblocks and Challenges

There has been significant progress in recent years with direct cardiac reprogramming through important discoveries in understanding the mechanism of reprogramming and the biology of cardiac development. However, several challenges must be addressed prior to clinical translation of this technology. The reprogramming efficiency must be increased in order to generate enough cells *in vitro* for transplantation. One avenue that has the potential to generate the number of cells needed for transplantation is reprogramming to CPCs, which can be expanded *in vitro* before transplantation. The retention, integration and maturation of iCMs or CPCs after transplantation remains a concern. Multiplex immunostaining and patch clamp analysis has also revealed the presence of all three cardiomyocyte types (atrial, ventricular and

pacemaker) in iCMs therefore increasing the risk of arrhythmias [16]. There is a need to develop techniques to generate specific subtypes of cardiomyocytes for both *in vitro* and *in vivo* direct reprogramming. A safe and effective approach to delivering and targeting reprogramming factors *in vivo* will be needed to circumvent *in vitro* reprogramming completely.

Transcription factors, inhibitors, cytokines, non-coding RNAs, and epigenetic modulators have been shown to be important for direct cardiac reprogramming. However, studies have uncovered variable reproducibility between different labs, leading to wide differences in reprogramming efficiency, maturity, and characteristics of the iCMs. These inconsistencies can be attributed to many factors other than the reprogramming factors themselves. First, the components of culture media(s) used during reprogramming widely varies from group to group along with the duration of reprogramming before analysis. Additionally, the induction time, the type of fibroblasts, the amount and sequence of factors used along with the time exposed to reprogramming factors are different between protocols. Moreover, the criteria used to measure the outcome and success are inconsistent and not standardized in the field. Reprogramming success is measured by some as the presence of cardiac-related structural proteins on immunostaining, while others employ a much more detailed approach including appearance of spontaneous beating along with gene and protein expression data. Even differences in the cardiac markers used to analyze the reprogramming efficiency, cardiac troponin T (cTnT) vs alpha myosin heavy chain (αMHC) vs GCaMP activity etc., and the method of measurement, flow cytometry vs immunofluorescence (IF), make comparisons among studies difficult. The myriad of criteria and stringency that have been used to evaluate

reprogramming efficiency have been summarized by Addis and Epstein and are presented in Table 6 [64].

Table 6: Criteria to evaluate reprogramming efficiency. Adapted from Addis and Epstein [64].

Characteristic	Stringency	Assay technique(s)
Gene expression	Low	RT-qPCR
		Reporter transgene
Protein expression	Low	Immunostaining
		Flow cytometry
		Western Blot
Transcriptome and epigenetic analysis	High	Microarray
		RNA-seq
		ChIP-seq
		ATAC-seq
Contraction and force generation	High	Spontaneous
		Chemical stimulation
		Electrical stimulation
		Three dimensional bioengineered platforms
Electrophysiological	High	Patch Clamp
		Microelectrode arrays
		Optical Mapping
Calcium transients and electrical coupling	High	Calcium sensitive dyes
		Genetically encoded indicators (GCaMP)
		Optical Mapping
Functional improvement	High	Echocardiography

Optimization of the minimal yet sufficient combination of factors to improve reprogramming requires further research. Studies presented here have revealed that simply expressing a few core transcription factors is not sufficient for efficient cardiac

reprogramming [65]. There may also be a dosage and temporal requirement for reprogramming factors [60, 61]. Other factors such as activated cellular signaling processes and epigenetic landscape should be considered to improve efficiency and quality. For example, when the TGF-β signaling pathway was disrupted by small molecules or when important cardiac regulatory miRNAs were added, an increase in reprogramming was observed, supporting the hypothesis that a multi-faceted approach is likely necessary to achieve high reprogramming efficiency. Furthermore, these studies highlight the significant differences between mouse and human reprogramming as well as the effect of the starting fibroblast type (MEFs, tail-tip or cardiac fibroblasts).

Are All Fibroblasts Created Equal?

Cardiac fibroblasts are ideal targets for direct reprogramming as they are the most prominent cell type within the heart and play key roles in regulating normal myocardial function as well as adverse remodeling following injury. Various mouse and/or human fibroblast sources have been tested, including mouse embryonic fibroblasts, tail-tip fibroblasts, and dermal fibroblasts, with varying results, suggesting the importance of the starting cell type for direct reprogramming. It is also interesting to note that *in vivo* reprogramming has been reported to be more efficient than *in vitro*, despite that fact that upon injury, cardiac fibroblast express TGF-β, which has been shown to be inhibitory to reprogramming *in vitro*. Furthermore, our lab has shown that cardiac fibroblasts are a heterogenous population from different embryonic origins [66]. It is possible to postulate that perhaps a subpopulation of cardiac fibroblasts may be more susceptible to reprogramming depending on their developmental origin. Further understanding of the epigenetic landscape of fibroblasts and their susceptibility to direct reprogramming would

be of great use to the field. This would also open up the possibility for repairing the heart by targeting specific fibroblast populations.

Conclusion

In this review, we discussed the reprogramming of fibroblasts into cardiomyocyte-like cells using transcription factors, small molecules, miRNAs, and other biologics for the treatment of heart failure (Figure 20). Despite the current limitations that exist with direct cardiac reprogramming, this technology offers great promise for cardiac regeneration therapy. It is clear that the reprogramming process is very complex and that many factors have profound influence over this process. Continued research of key transcription factors, non-coding RNAs, small molecules, reprogramming mechanisms, delivery and targeting methods and biomaterials will help advance direct cardiac reprogramming to large animal models and ultimately for the treatment of heart failure.

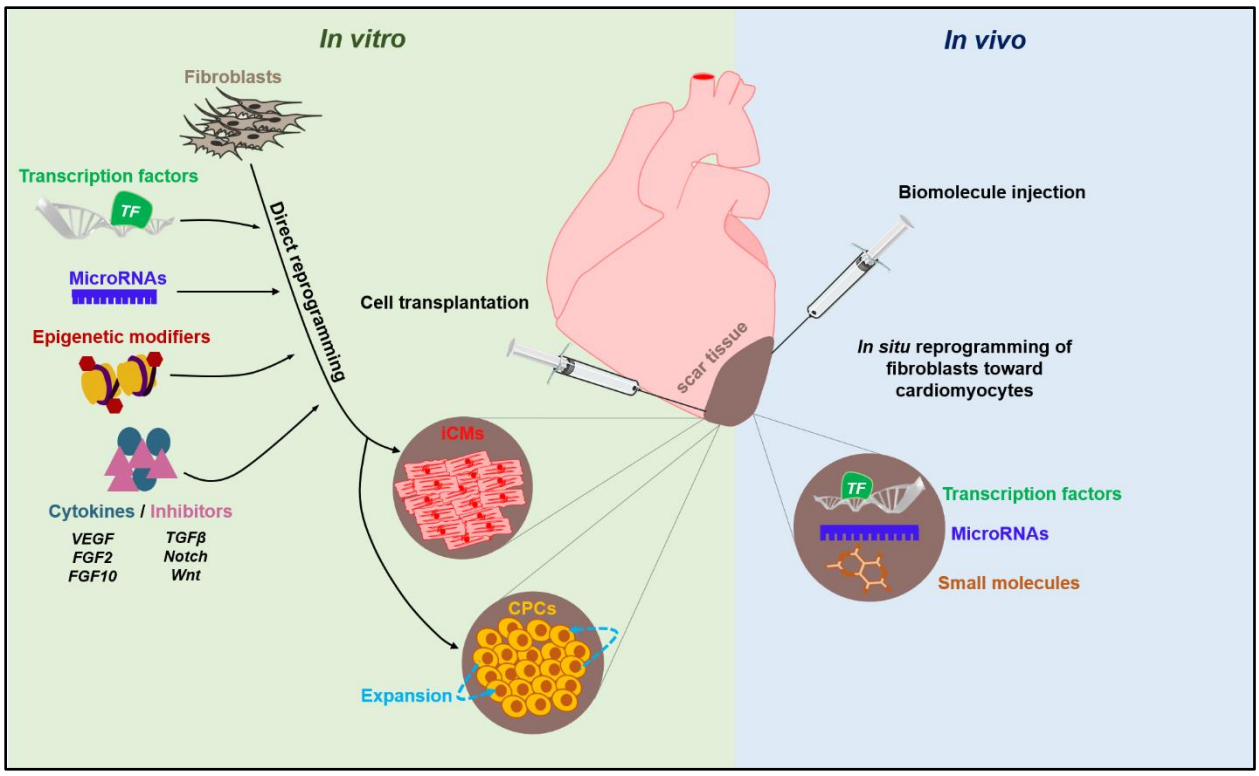


Figure 20. Schematic showing the current and future applications of direct cardiac reprogramming.

References

1. Braunwald E. The war against heart failure: the Lancet lecture. *Lancet*. 2015;385:812-24.

2. Heidenreich PA, Albert NM, Allen LA, Bluemke DA, Butler J, Fonarow GC, Ikonomidis JS, Khavjou O, Konstam MA, Maddox TM, Nichol G, Pham M, Piña IL, Trogdon JG, Committee AHAAC, Council on Arteriosclerosis TraVB, Intervention CoCRa, Cardiology CoC, Prevention CoEa and Council S. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. *Circ Heart Fail*. 2013;6:606-19.

3. Cahill TJ, Choudhury RP and Riley PR. Heart regeneration and repair after myocardial infarction: translational opportunities for novel therapeutics. *Nat Rev Drug Discov*. 2017.

4. Taylor CJ, Roalfe AK, Iles R and Hobbs FD. Ten-year prognosis of heart failure in the community: follow-up data from the Echocardiographic Heart of England Screening (ECHOES) study. *Eur J Heart Fail*. 2012;14:176-84.

5. Davis RL, Weintraub H and Lassar AB. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell*. 1987;51:987-1000.

6. Choi J, Costa ML, Mermelstein CS, Chagas C, Holtzer S and Holtzer H. MyoD converts primary dermal fibroblasts, chondroblasts, smooth muscle, and retinal pigmented epithelial cells into striated mononucleated myoblasts and multinucleated myotubes. *Proc Natl Acad Sci U S A*. 1990;87:7988-92.

7. Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Südhof TC and Wernig M. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature*. 2010;463:1035-41.

8. Yu B, He ZY, You P, Han QW, Xiang D, Chen F, Wang MJ, Liu CC, Lin XW, Borjigin U, Zi XY, Li JX, Zhu HY, Li WL, Han CS, Wangensteen KJ, Shi Y, Hui LJ, Wang X and Hu YP. Reprogramming fibroblasts into bipotential hepatic stem cells by defined factors. *Cell Stem Cell*. 2013;13:328-40.

9. Huang P, He Z, Ji S, Sun H, Xiang D, Liu C, Hu Y, Wang X and Hui L. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature*. 2011;475:386-9.

10. Xie H, Ye M, Feng R and Graf T. Stepwise reprogramming of B cells into macrophages. *Cell*. 2004;117:663-76.

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

12. Hiroi Y, Kudoh S, Monzen K, Ikeda Y, Yazaki Y, Nagai R and Komuro I. Tbx5 associates with Nkx2-5 and synergistically promotes cardiomyocyte differentiation. *Nat Genet*. 2001;28:276-80.

13. Perrino C and Rockman HA. GATA4 and the two sides of gene expression reprogramming. *Circ Res*. 2006;98:715-6.

14. Dodou E, Verzi MP, Anderson JP, Xu SM and Black BL. Mef2c is a direct transcriptional target of ISL1 and GATA factors in the anterior heart field during mouse embryonic development. *Development*. 2004;131:3931-42.

15. Dai YS, Cserjesi P, Markham BE and Molkentin JD. The transcription factors GATA4 and dHAND physically interact to synergistically activate cardiac gene expression through a p300-dependent mechanism. *J Biol Chem*. 2002;277:24390-8.

16. Nam YJ, Lubczyk C, Bhakta M, Zang T, Fernandez-Perez A, McAnally J, Bassel-Duby R, Olson EN and Munshi NV. Induction of diverse cardiac cell types by reprogramming fibroblasts with cardiac transcription factors. *Development*. 2014;141:4267-78.

17. Hirai H, Katoku-Kikyo N, Keirstead SA and Kikyo N. Accelerated direct reprogramming of fibroblasts into cardiomyocyte-like cells with the MyoD transactivation domain. *Cardiovasc Res*. 2013;100:105-13.

18. Addis RC, Ifkovits JL, Pinto F, Kellam LD, Esteso P, Rentschler S, Christoforou N, Epstein JA and Gearhart JD. Optimization of direct fibroblast reprogramming to cardiomyocytes using calcium activity as a functional measure of success. *J Mol Cell Cardiol*. 2013;60:97-106.

19. Protze S, Khattak S, Poulet C, Lindemann D, Tanaka EM and Ravens U. A new approach to transcription factor screening for reprogramming of fibroblasts to cardiomyocyte-like cells. *J Mol Cell Cardiol*. 2012;53:323-32.

20. Christoforou N, Chellappan M, Adler AF, Kirkton RD, Wu T, Addis RC, Bursac N and Leong KW. Transcription factors MYOCD, SRF, Mesp1 and SMARCD3 enhance the cardio-inducing effect of GATA4, TBX5, and MEF2C during direct cellular reprogramming. *PLoS One*. 2013;8:e63577.

21. Islas JF, Liu Y, Weng KC, Robertson MJ, Zhang S, Prejusa A, Harger J, Tikhomirova D, Chopra M, Iyer D, Mercola M, Oshima RG, Willerson JT, Potaman VN and Schwartz RJ. Transcription factors ETS2 and MESP1 transdifferentiate human

dermal fibroblasts into cardiac progenitors. *Proc Natl Acad Sci U S A*. 2012;109:13016-21.

22. Wada R, Muraoka N, Inagawa K, Yamakawa H, Miyamoto K, Sadahiro T, Umei T, Kaneda R, Suzuki T, Kamiya K, Tohyama S, Yuasa S, Kokaji K, Aeba R, Yozu R, Yamagishi H, Kitamura T, Fukuda K and Ieda M. Induction of human cardiomyocyte-like cells from fibroblasts by defined factors. *Proc Natl Acad Sci U S A*. 2013;110:12667-72.

23. Fu JD, Stone NR, Liu L, Spencer CI, Qian L, Hayashi Y, Delgado-Olguin P, Ding S, Bruneau BG and Srivastava D. Direct reprogramming of human fibroblasts toward a cardiomyocyte-like state. *Stem Cell Reports*. 2013;1:235-47.

24. Ao A, Hao J, Hopkins CR and Hong CC. DMH1, a novel BMP small molecule inhibitor, increases cardiomyocyte progenitors and promotes cardiac differentiation in mouse embryonic stem cells. *PLoS One*. 2012;7:e41627.

25. Cai W, Guzzo RM, Wei K, Willems E, Davidovics H and Mercola M. A Nodal-to-TGFβ cascade exerts biphasic control over cardiopoiesis. *Circ Res*. 2012;111:876-81.

26. Ifkovits JL, Addis RC, Epstein JA and Gearhart JD. Inhibition of TGFβ signaling increases direct conversion of fibroblasts to induced cardiomyocytes. *PLoS One*. 2014;9:e89678.

27. Mohamed TM, Stone NR, Berry EC, Radzinsky E, Huang Y, Pratt K, Ang YS, Yu P, Wang H, Tang S, Magnitsky S, Ding S, Ivey KN and Srivastava D. Chemical Enhancement of In Vitro and In Vivo Direct Cardiac Reprogramming. *Circulation*. 2017;135:978-995.

28. Abad M, Hashimoto H, Zhou H, Morales MG, Chen B, Bassel-Duby R and Olson EN. Notch Inhibition Enhances Cardiac Reprogramming by Increasing MEF2C Transcriptional Activity. *Stem Cell Reports*. 2017;8:548-560.

29. Zhao Y, Londono P, Cao Y, Sharpe EJ, Proenza C, O'Rourke R, Jones KL, Jeong MY, Walker LA, Buttrick PM, McKinsey TA and Song K. High-efficiency reprogramming of fibroblasts into cardiomyocytes requires suppression of pro-fibrotic signalling. *Nat Commun*. 2015;6:8243.

30. Jayawardena TM, Egemnazarov B, Finch EA, Zhang L, Payne JA, Pandya K, Zhang Z, Rosenberg P, Mirotsou M and Dzau VJ. MicroRNA-mediated in vitro and in vivo direct reprogramming of cardiac fibroblasts to cardiomyocytes. *Circ Res*. 2012;110:1465-73.

31. Zhou H, Dickson ME, Kim MS, Bassel-Duby R and Olson EN. Akt1/protein kinase B enhances transcriptional reprogramming of fibroblasts to functional cardiomyocytes. *Proc Natl Acad Sci U S A*. 2015;112:11864-9.

32. Yamakawa H, Muraoka N, Miyamoto K, Sadahiro T, Isomi M, Haginiwa S, Kojima H, Umei T, Akiyama M, Kuishi Y, Kurokawa J, Furukawa T, Fukuda K and Ieda M. Fibroblast Growth Factors and Vascular Endothelial Growth Factor Promote Cardiac Reprogramming under Defined Conditions. *Stem Cell Reports*. 2015;5:1128-42.

33. Rao PK, Toyama Y, Chiang HR, Gupta S, Bauer M, Medvid R, Reinhardt F, Liao R, Krieger M, Jaenisch R, Lodish HF and Blelloch R. Loss of cardiac microRNA-mediated regulation leads to dilated cardiomyopathy and heart failure. *Circ Res*. 2009;105:585-94.

34. Liu N, Bezprozvannaya S, Williams AH, Qi X, Richardson JA, Bassel-Duby R and Olson EN. microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. *Genes Dev*. 2008;22:3242-54.

35. Muraoka N, Yamakawa H, Miyamoto K, Sadahiro T, Umei T, Isomi M, Nakashima H, Akiyama M, Wada R, Inagawa K, Nishiyama T, Kaneda R, Fukuda T, Takeda S, Tohyama S, Hashimoto H, Kawamura Y, Goshima N, Aeba R, Yamagishi H, Fukuda K and Ieda M. MiR-133 promotes cardiac reprogramming by directly repressing Snai1 and silencing fibroblast signatures. *EMBO J*. 2014;33:1565-81.

36. Jayawardena TM, Finch EA, Zhang L, Zhang H, Hodgkinson CP, Pratt RE, Rosenberg PB, Mirotsou M and Dzau VJ. MicroRNA induced cardiac reprogramming in vivo: evidence for mature cardiac myocytes and improved cardiac function. *Circ Res*. 2015;116:418-24.

37. Nam YJ, Song K, Luo X, Daniel E, Lambeth K, West K, Hill JA, DiMaio JM, Baker LA, Bassel-Duby R and Olson EN. Reprogramming of human fibroblasts toward a cardiac fate. *Proc Natl Acad Sci U S A*. 2013;110:5588-93.

38. Eulalio A, Mano M, Dal Ferro M, Zentilin L, Sinagra G, Zacchigna S and Giacca M. Functional screening identifies miRNAs inducing cardiac regeneration. *Nature*. 2012;492:376-81.

39. Singh VP, Mathison M, Patel V, Sanagasetti D, Gibson BW, Yang J and Rosengart TK. MiR-590 Promotes Transdifferentiation of Porcine and Human Fibroblasts Toward a Cardiomyocyte-Like Fate by Directly Repressing Specificity Protein 1. *J Am Heart Assoc*. 2016;5.

40. Johnsson P, Lipovich L, Grandér D and Morris KV. Evolutionary conservation of long non-coding RNAs; sequence, structure, function. *Biochim Biophys Acta*. 2014;1840:1063-71.

41. Gomes CPC, Spencer H, Ford KL, Michel LYM, Baker AH, Emanueli C, Balligand JL, Devaux Y and network C. The Function and Therapeutic Potential of Long Non-coding RNAs in Cardiovascular Development and Disease. *Mol Ther Nucleic Acids*. 2017;8:494-507.

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

43. Xue Z, Hennelly S, Doyle B, Gulati AA, Novikova IV, Sanbonmatsu KY and Boyer LA. A G-Rich Motif in the lncRNA Braveheart Interacts with a Zinc-Finger Transcription Factor to Specify the Cardiovascular Lineage. *Mol Cell*. 2016;64:37-50.

44. Grote P, Wittler L, Hendrix D, Koch F, Währisch S, Beisaw A, Macura K, Bläss G, Kellis M, Werber M and Herrmann BG. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Dev Cell*. 2013;24:206-14.

45. Wang Z, Zhang XJ, Ji YX, Zhang P, Deng KQ, Gong J, Ren S, Wang X, Chen I, Wang H, Gao C, Yokota T, Ang YS, Li S, Cass A, Vondriska TM, Li G, Deb A, Srivastava D, Yang HT, Xiao X, Li H and Wang Y. The long noncoding RNA Chaer defines an epigenetic checkpoint in cardiac hypertrophy. *Nat Med*. 2016;22:1131-1139.

46. Wang Z and Wang Y. Dawn of the Epi-LncRNAs: new path from Myheart. *Circ Res*. 2015;116:235-6.

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

48. Zhou Y, Wang L, Vaseghi HR, Liu Z, Lu R, Alimohamadi S, Yin C, Fu JD, Wang GG, Liu J and Qian L. Bmi1 Is a Key Epigenetic Barrier to Direct Cardiac Reprogramming. *Cell Stem Cell*. 2016;18:382-95.

49. Liu Z, Wang L, Welch JD, Ma H, Zhou Y, Vaseghi HR, Yu S, Wall JB, Alimohamadi S, Zheng M, Yin C, Shen W, Prins JF, Liu J and Qian L. Single-cell transcriptomics reconstructs fate conversion from fibroblast to cardiomyocyte. *Nature*. 2017;551:100-104.

50. Warren L, Manos PD, Ahfeldt T, Loh YH, Li H, Lau F, Ebina W, Mandal PK, Smith ZD, Meissner A, Daley GQ, Brack AS, Collins JJ, Cowan C, Schlaeger TM and Rossi DJ. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell*. 2010;7:618-30.

51. Mandal PK and Rossi DJ. Reprogramming human fibroblasts to pluripotency using modified mRNA. *Nat Protoc*. 2013;8:568-82.

52. Durkin ME, Qian X, Popescu NC and Lowy DR. Isolation of Mouse Embryo Fibroblasts. *Bio Protoc*. 2013;3.

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

54. Lalit PA, Salick MR, Nelson DO, Squirrell JM, Shafer CM, Patel NG, Saeed I, Schmuck EG, Markandeya YS, Wong R, Lea MR, Eliceiri KW, Hacker TA, Crone WC, Kyba M, Garry DJ, Stewart R, Thomson JA, Downs KM, Lyons GE and Kamp TJ. Lineage Reprogramming of Fibroblasts into Proliferative Induced Cardiac Progenitor Cells by Defined Factors. *Cell Stem Cell*. 2016;18:354-67.

55. Lalit PA, Rodriguez AM, Downs KM and Kamp TJ. Generation of multipotent induced cardiac progenitor cells from mouse fibroblasts and potency testing in ex vivo mouse embryos. *Nat Protoc*. 2017;12:1029-1054.

56. Zhang Y, Cao N, Huang Y, Spencer CI, Fu JD, Yu C, Liu K, Nie B, Xu T, Li K, Xu S, Bruneau BG, Srivastava D and Ding S. Expandable Cardiovascular Progenitor Cells Reprogrammed from Fibroblasts. *Cell Stem Cell*. 2016;18:368-81.

57. Qian L, Huang Y, Spencer CI, Foley A, Vedantham V, Liu L, Conway SJ, Fu JD and Srivastava D. In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. *Nature*. 2012;485:593-8.

58. Ma H, Wang L, Yin C, Liu J and Qian L. In vivo cardiac reprogramming using an optimal single polycistronic construct. *Cardiovasc Res*. 2015;108:217-9.

59. Mathison M, Singh VP, Gersch RP, Ramirez MO, Cooney A, Kaminsky SM, Chiuchiolo MJ, Nasser A, Yang J, Crystal RG and Rosengart TK. "Triplet" polycistronic vectors encoding Gata4, Mef2c, and Tbx5 enhances postinfarct ventricular functional improvement compared with singlet vectors. *J Thorac Cardiovasc Surg*. 2014;148:1656-1664.e2.

60. Wang L, Liu Z, Yin C, Asfour H, Chen O, Li Y, Bursac N, Liu J and Qian L. Stoichiometry of Gata4, Mef2c, and Tbx5 influences the efficiency and quality of induced cardiac myocyte reprogramming. *Circ Res*. 2015;116:237-44.

61. Umei TC, Yamakawa H, Muraoka N, Sadahiro T, Isomi M, Haginiwa S, Kojima H, Kurotsu S, Tamura F, Osakabe R, Tani H, Nara K, Miyoshi H, Fukuda K and Ieda M. Single-Construct Polycistronic Doxycycline-Inducible Vectors Improve Direct Cardiac Reprogramming and Can Be Used to Identify the Critical Timing of Transgene Expression. *Int J Mol Sci*. 2017;18.

62. Fu Y, Huang C, Xu X, Gu H, Ye Y, Jiang C, Qiu Z and Xie X. Direct reprogramming of mouse fibroblasts into cardiomyocytes with chemical cocktails. *Cell Res*. 2015;25:1013-24.

63. Cao N, Huang Y, Zheng J, Spencer CI, Zhang Y, Fu JD, Nie B, Xie M, Zhang M, Wang H, Ma T, Xu T, Shi G, Srivastava D and Ding S. Conversion of human fibroblasts into functional cardiomyocytes by small molecules. *Science*. 2016;352:1216-20.

64. Addis RC and Epstein JA. Induced regeneration--the progress and promise of direct reprogramming for heart repair. *Nat Med*. 2013;19:829-36.

65. Chen JX, Krane M, Deutsch MA, Wang L, Rav-Acha M, Gregoire S, Engels MC, Rajarajan K, Karra R, Abel ED, Wu JC, Milan D and Wu SM. Inefficient reprogramming of fibroblasts into cardiomyocytes using Gata4, Mef2c, and Tbx5. *Circ Res*. 2012;111:50-5.

66. Ali SR, Ranjbarvaziri S, Talkhabi M, Zhao P, Subat A, Hojjat A, Kamran P, Müller AM, Volz KS, Tang Z, Red-Horse K and Ardehali R. Developmental heterogeneity of cardiac fibroblasts does not predict pathological proliferation and activation. *Circ Res*. 2014;115:625-35.