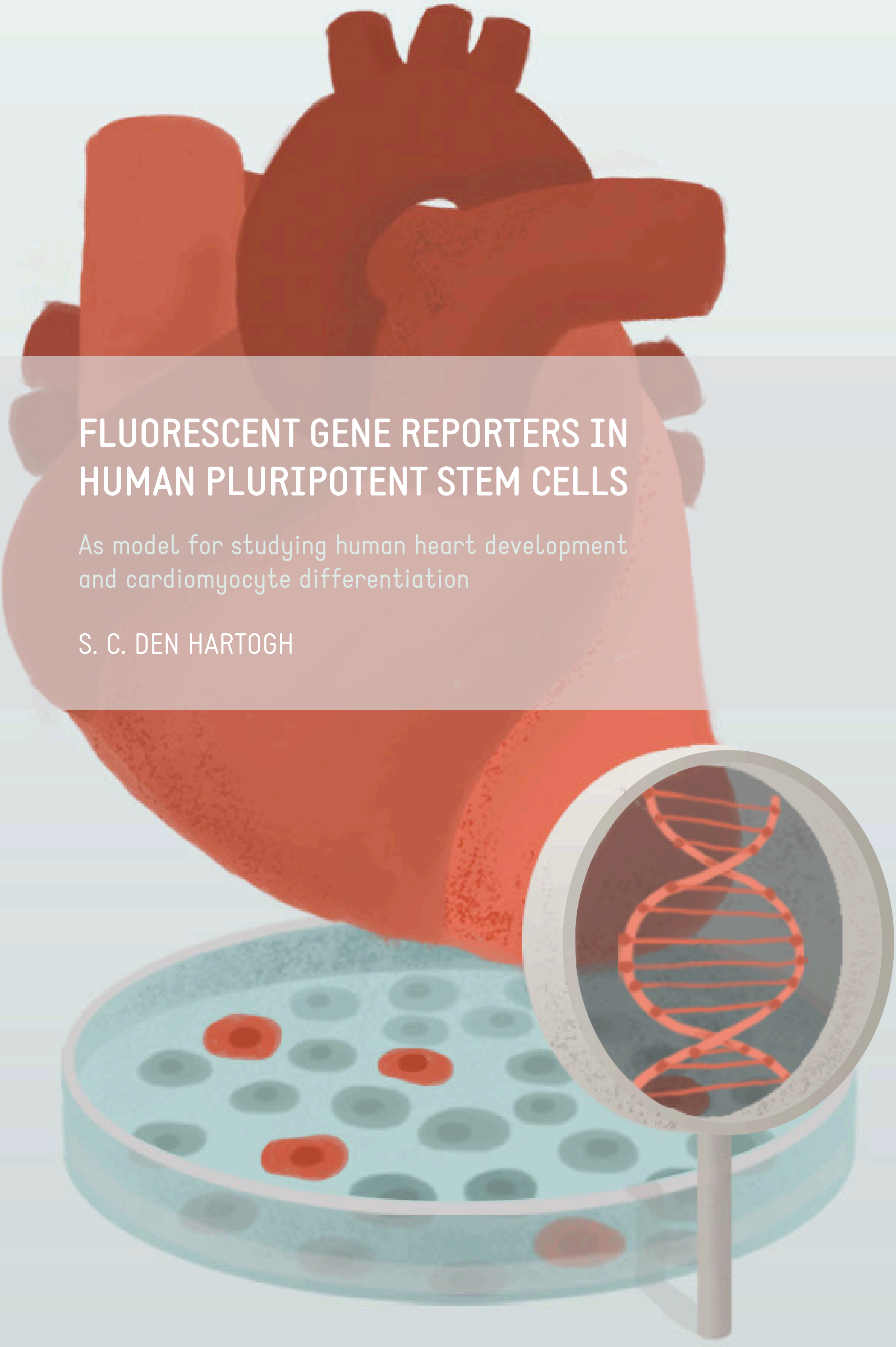
Author: Hartogh, S.C. den
Title: Fluorescent gene reporters in human pluripotent stem cells : as model for studying human heart development and cardiomyocyte differentiation
Issue Date: 2016-05-18



FLUORESCENT GENE REPORTERS IN HUMAN PLURIPOTENT STEM CELLS

As model for studying human heart development and cardiomyocyte differentiation

S. C. DEN HARTOGH

FLUORESCENT GENE REPORTERS IN HUMAN PLURIPOTENT STEM CELLS

As model for studying human heart development and cardiomyocyte differentiation

Sabine Charlotte den Hartogh

Promotoren: Prof. dr. C. L. Mummery, Prof. dr. R. Passier

Promotiecommissie:
Dr. L. van Laake (UMC Utrecht)
Prof. Dr. R. Zweigerdt (Hannover Medical Centre)
Prof. Dr. F. Staal
Prof. Dr. M.J. Goumans
Dr. E. van Rooij (Hubrecht Institute)

Fluorescent gene reporters in human pluripotent stem cells
PhD thesis, Leiden University

The research described in this thesis was supported by a grant of Rembrandt Institute of Cardiovascular Science.

Financial support by the Dutch Heart Foundation for the publication of this thesis is gratefully acknowledged.

Het verschijnen van dit proefschrift werd mede mogelijk gemaakt door de steun van Stichting Proefdiervrij.

ISBN: 9789462954823
Lay-Out: Lonneke Weuring
Printing: Proefschriftmaken.nl ||Uitgeverij BOXPress

FLUORESCENT GENE REPORTERS IN HUMAN PLURIPOTENT STEM CELLS

As model for studying human heart development and cardiomyocyte differentiation

Proefschrift

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden, op gezag van
Rector Magnificus prof.mr. C.J.J.M. Stolker
volgens besluit van het College voor Promoties, te verdedigen
op woensdag 18 mei 2016 klokke 16:15 uur

door

Sabine Charlotte den Hartogh
geboren op 16 juli 1987, te Nijkerk

Aan mijn ouders
Voor de wetenschap

CONTENTS

Chapter 1 06	**General Introduction**
Chapter 2 22	**Fluorescent reporters in human luripotent stem cells: contributions to cardiac differentiation and their applications in cardiac disease and toxicity** Den Hartogh SC, Passier R. Stem Cells. 2015 Oct 8. doi: 10.1002/stem.2196. In Press
Chapter 3 52	**Dual Reporter MESP1$^{mcherry/W}$NKX2-5$^{eGFP/W}$ hESCs Enable Studying Early Human Cardiac Differentiation** Den Hartogh SC, Schreurs C, Monshouwer-Kloots JJ, Davis RP, Elliott DA, Mummery CL, Passier R. Stem Cells. 2015 Jan;33(1):56-67. doi: 10.1002/stem.1842.
Chapter 4 82	**A comprehensive gene expression analysis at sequential stages of in vitro cardiac differentiation from isolated MESP1-expressing-mesoderm progenitors** Sabine den Hartogh, Katy Wolstencroft, Christine Mummery, Robert Passier. Scientific Reports 6, (2016) doi:10.1038/srep19386
Chapter 5 112	**Culturing Cardiac Conduction System Cells from Human Pluripotent Stem Cells using TBX3 regulatory elements: Technical and Biological Challenges**
Chapter 6 128	**A Novel Approach for the Pure Isolation of Ventricular Cardiomyocytes from human Embryonic Stem Cells: a dual MYL2-T2A-mcherry NKX2-5-eGFP Fluorescent Reporter Line**
Chapter 7 142	**Analysis of Pre-Cardiac MESP1 Progenitor Proliferation at Single Cell Level using an eGFP-Anillin Reporter Line** Sabine den Hartogh, Lu Cao, Fons Verbeek, Robert Passier. *Parts of this chapter are submitted for publication (2016)*
Chapter 8 170	**General Discussion**
185	**Summary / Samenvatting**
192	**Dankwoord / Acknowledgements**
194	**Curriculum Vitae**
195	**List of Publications**

CHAPTER 1

Introduction

1.1 Pluripotent Stem Cell-derived Cardiomyocytes: Opportunities and Translational Applications.

The derivation of human embryonic stem cells (hESCs) (1998)[1] and ground-breaking discovery of human induced pluripotent stem cells (hiPSCs) (2007)[2] initiated an era with exciting opportunities for the development of cell-based therapies, disease models and in vitro drug discovery and toxicity screening. In addition, human pluripotent stem cells (hPSCs) offer advantages over conventional methods for studying human biological and developmental processes. HESCs are derived from the inner cell mass of blastocyst stage embryos and proliferate indefinitely in vitro, while maintaining their undifferentiated state when cultured under appropriate conditions. hiPSCs have a similar indefinite growth capacity, but are of less ethical concern as they can be derived through the reprogramming of a variety of human somatic cells, such as skin fibroblasts, blood cells, teeth, or even epithelial cells from urine samples[3-5]. Both hESCs and hiPSCs are hPSCs and have the potential to differentiate into derivatives of all three primary germ layers: ectoderm, endoderm, and mesoderm [6, 7]. The heart, skeletal muscle, mesenteries, kidney, ovary, and testis are mostly derived from mesoderm so that deriving these cell types requires first that mesoderm is formed **(Fig. 1).** The ability to differentiate hPSCs in vitro to any cell-type of the human body, including cardiomyocytes, prompted researchers to investigate new translational opportunities, such as cellular therapy of the damaged heart **(Fig. 2)**[8].

Cellular Therapy

Ischemic heart disease is one of the leading causes of death worldwide and is increasing due to population ageing[9]. Because postnatal cardiomyocytes have little or no regenerative capacity[10], the adult heart is unable to restore the large numbers of cardiomyocytes lost after infarct, resulting in adaptation mechanisms, such as scar formation and cardiomyocyte hypertrophy, followed by decrease in cardiac function. This was the basis of the initial interest in deriving cardiomyocytes from hESCs. However, many obstacles have been encountered during transplantation of hPSC cardiomyocytes in both rodents and primates[8, 11]. These challenges included the method of cell delivery, the cell type to be injected, the risk of arrhythmias induced by engrafted beating cardiomyocytes, electrical coupling between donor and host cells, immune rejection, poor donor cell survival and the large number of cells required for complete cell replacement. These challenges contributed to failure to achieve long-term cardiac functional improvements[8, 12].

Drug Discovery and Development

An alternative opportunity that emerged is the use of hPSC-derived CMs in human drug discovery and safety assessment, as well as modelling human cardiac disease [13, 14]. Unanticipated drug-induced cardiotoxicity is a major problem with life-threatening consequences; all newly developed drugs need to be screened for potential cardiotoxic effects before they reach the market. Since human primary cardiomyocyte cultures are difficult to obtain and maintain in culture, other alternatives, such as Chinese hamster ovary (CHO) cells, human embryonic kidney cells overexpressing the hERG channel, or in vivo animal models are currently used by pharmaceutical industry[15]. Interspecies differences related to heart function and characterization, such as heart rate, ion currents, and expression of genes encoding signaling, membrane and structural proteins[16], point out the importance of developing human-based models for drug discovery and disease modeling. Human PSC-CMs represent a promising in vitro model since they resemble native human cardiomyocytes[17]. hPSC-CMs have been shown to respond to clinically relevant drugs and are able to predict cardiotoxic and pro-arrythmic side effects[18]. hPSC-CMs-based assay technology is gaining interest from the pharmaceutical industry, and a variety of screening assays are currently being developed at scalable levels for cost-effective and efficient high throughput level compound screenings. Moreover, more complex assays, such as co-culture or 3D- tissue engineered assays, which may even more faithfully recapitulate cardiac physiology and function, are underway [19, 20].

Developmental Biology and Cell Differentiation

Besides the applicability of hPSC-derived CMs for in vivo and in vitro translational purposes, another prospect that emerged with the realization of in vitro hPSC-CMs cultures is their use as in vitro model for studying human heart development and the onset of congenital heart disease, explained in more detail in the section below. Vice versa, a better understanding of molecular pathways involved in the formation of the developing heart, may contribute to the development and further improvements of protocols that generate efficient directed differentiations of hPSCs towards specific cell-types of the heart. Clearly, this is advantageous for therapeutic and pharmaceutical applications, including assay development on the field of drug discovery and disease modelling.

1.2 Heart Development: A Blueprint for in vitro Cardiomyocyte Differentiation

Anatomy

The heart is the first organ to form during human embryonic development, at around the third week of gestation[21]. Our current detailed understanding of early heart development has been based on numerous studies in model organisms including rodents **(Fig. 3)**, zebrafish, chick, and amphibian and a limited number of studies in the early human embryo[22-24]. During human development, prior to gastrulation, the bilaminar embryonic disc is almost flat and consists of two layers: the epiblast and hypoblast. The epiblast will give rise to the embryo and extra-embryonic structures, while the hypoblast will contribute to the extra-embryonic membranes. Upon early gastrulation, epiblast cells ingress through the primitive node and the primitive streak to form the ectodermal layer, and undergoes epithelial-to-mesenchymal transition (EMT) to form the mesodermal and endodermal layer

(Fig. 4a). The mesodermal population, marked by the expression of key transcription factor Brachyury T arise from the posterior mesoderm, and migrate to the anterior-lateral regions of the embryo, at both sides of the midline. This is followed by formation of the earliest mesodermal pre-cardiac progenitor cells, marked by expressing of key transcription factor MESP1. Myocardial markers are first detected at this stage, where two molecularly distinct progenitor populations are identified, referred to the First Heart Field (FHF) and the Second Heart Field (SHF). The cells then extend across the midline to form a crescent-shaped epithelium, the cardiac crescent. The FHF differentiates within the cardiac crescent, fuses at the midline, and forms the early heart tube, responsible for the distribution of nutrients and oxygen in the embryo[25, 26]. The SHF is widely marked by LIM homeobox transcription factor 1 (ISL1), and is located medially to the cardiac crescent and at later stages behind the forming heart tube. The SHF, contributes to the venous as well as the arterial pole of the heart tube. The tube undergoes rightward looping, with its posterior region moving anteriorly. Derivatives from the FHF form the entire left ventricle as well as other parts of the heart, where SHF derivatives contribute to the right ventricle and atria, and the entire outflow tract. The heart is shaped by the looping process, expansion of the myocardium, which leads to the formation of the cardiac chambers, septation and further maturation.

Cardiovascular Progenitor Lineages

As mentioned above, the mesoderm transcription factor MESP1 is one of the earliest key (pre-) cardiac transcription factors, activating the expression of downstream cardiac transcription factors, including NKX2-5, GATA-4, HAND2, MYOCARDIN, and MEF2C[27]. Although MESP1 progenitors have a broader contribution to the embryo than only the heart[28-30], lineage tracing in mice demonstrated that almost all the cells of the heart are derived from MESP1 progenitors[30, 31] **(Fig. 4b)**. A more specified subset of MESP1 progenitor derivatives was recently suggested by their expression of Smarcd3[32]. However, also from these progenitors, extra-embryonic derivatives outside the heart could be identified. The earliest molecularly distinct cardiac progenitor populations are the FHF and the SHF, of which the latter is marked by enriched Isl1 expression[25, 32-34]. The earliest transcription factors that start to be expressed in the cardiac crescent are NKX2-5 and GATA-4, both identified as downstream target genes of transcription factor MESP1[35].

Transcription factor NKX2-5, the vertebrate *tinman* homologue of the Drosophila fruitfly, is essential for heart formation and marks definitive multipotent cardiac progenitors. The multi-chambered septated heart consists of several subtypes of cardiac cells, including working myocardium cardiomyocytes (atrial and ventricular cardiomyocytes), smooth muscle cells, endothelial cells, fibroblasts and cardiomyocytes from the cardiac conduction system (CCS), consisting of the sinoatrial node (SAN), the atrioventricular node (AVN), bundle of His and Purkinje fibers.

Molecular Signalling

The endoderm germ layer plays an important role for the formation of the heart by exerting both stimulatory as well as inhibitory signals. Early cardiogenesis is modulated by four major signalling pathways of which the ligands originate in the endoderm: bone morphogenetic protein (BMP), nodal/activin, Wnt/β-catenin, and fibroblast growth factor (FGF). Both the BMP and WNT/β-catenin pathways have a biphasic role during cardiogenesis. BMPs and Wnts, are required for mesoderm specification[36]. In contrast, subsequent cardiac progenitor differentiation and activation of key myocardial regulatory genes, such as Nkx2-5 and Gata-4, are repressed by BMP and WNT/β catenin signalling[37, 38]. Moreover, prior to their addition to the primary heart tube, SHF progenitors are exposed to increasing BMP and non-canonical Wnt signals, which are thought to be the primary drivers for myocardial differentiation[39]. During all phases of the developing heart, WNT ligands, Frizzled receptors, or extracellular WNT inhibitors are expressed in relevant tissues, with an aberrant role for high WNT signalling in mesoderm formation and CPC proliferation. Furthermore, WNT signalling also plays a prominent role in morphogenetic processes, valve formation, and establishment of the conduction system **(Fig. 5)**[38]. An in-depth knowledge of molecular signalling during early heart development is crucial for developing and optimizing protocols for the derivation of in vitro PSC-CMs and/or in vitro CPC proliferation.

In this thesis, we describe the use of fluorescent reporters in hPSC lines for a better understanding of the molecular mechanisms underlying early human cardiac lineage commitment and cardiac progenitor dynamics

1.3 Directed Differentiation of Pluripotent Stem Cells towards the Cardiac Lineage

In vitro Cardiac Differentiation Protocols

For defined, robust and controlled differentiation of specific cardiac lineages from hPSCs, lessons from early cardiac development have proven invaluable. They have supported the development of ways for hPSCs to provide an unlimited supply of cardiomyocytes from a single clonal source.

In recent years, various robust protocols have been established for efficient differentiation to cardiomyocytes from hPSCs[40-42]. The two principle approaches for differentiation are through aggregate formation or monolayer culture **(Fig. 6)**[41, 43]. For aggregate formation, dissociated hPSCs are seeded and aggregated by centrifugation in V-bottomed 96-well plates, leading to the formation of uniformly sized aggregates. The inclusion of polyvinyl alcohol is essential for initial aggregation of the cells. For simplicity, commercially available li-APEL or BPEL are often used[44]. Monolayer differentiation is carried out in culture media such as RPMI/B27. This is specially used for hESCs maintained in feeder-free culture systems, such as mTeSR or E8 (Stem Cell Technologies), as these do not easily form spin EBs

by centrifugation, but it is also applicable for hPSCs cultured on feeder layers.
Recently developed methods are based on serum-free culture conditions, involving the sequential activation or inhibition of signalling pathways involved in heart development: BMP4, Activin-A, and Wnt3a (or a GSK3 inhibitor, such as small molecule Chir99201) are added to monolayer hPSC cultures for 3 days, resulting in the temporal expression of MESP1+ progenitors, which are subsequently exposed to a Wnt signalling inhibitor, such as small molecule Xav939, to further differentiate to NKX2-5 expressing progenitors and cardiomyocytes **(Fig. 7)**[44-47]. Recently, Lian et al. and Burridge et al. published defined, growth factor–free monolayer protocols for differentiating feeder-free hPSCs to cardiomyocytes[41, 42]. These protocols involve the exposure of PSCs to GSK3 inhibitor Chir99201 for 1 or 2 days, followed by the treatment with Wnt inhibitor IWP2 for 2 days. These protocols yielded ±90% cTnT+ cells, with numbers of greater than 106 cardiomyocytes/cm2 in multiple hESC and iPSC lines .
hPSC-derived cardiomyocytes express sarcomeric proteins, such as cardiac Troponin T (cTnT), and cardiac transcription factors, including TBX5, TBX20, GATA4, MEF2C, and ISL1, chamber-specific proteins, including MLC2V and MLC2A, and ion channels (such as sodium, potassium, and calcium channels). To note, transcription factors mentioned above can also be expressed in tissue other than the heart. Therefore, defining the combinatorial expression of sets of transcription factors is always important in such studies.

In general, these cardiomyocyte differentiation protocols lead to a mixture of cardiomyocytes, consisting mostly of ventricular cardiomyocytes. In order to develop new therapeutic strategies treating diseases affecting specific cardiac cell subtypes, such as atrial or ventricular fibrillation or dysfunctions of the pacemaker and conduction tissues, or replacing ventricular cardiomyocytes following myocardial infarction, it is of importance to develop cardiomyocyte subtype-directed differentiation protocols[23, 48, 49].

In vitro PSC-derived Cardiac Progenitor Populations

hPSCs offer the opportunity to develop in vitro derived cardiac progenitor populations. These could be useful for understanding of how progenitor cells can be maintained, expanded, and differentiated to specific cardiac subtypes under defined and controlled differentiation conditions[39]. One such progenitor population is marked by the expression of MESP1, from which almost all cells of the heart are derived, as described before [30]. In a mouse study, using fluorescent labelling, ESC-derived pre-cardiac MESP1+ progenitors were isolated and characterized[50]. Here, they showed how these progenitors could be isolated based on Flk1+CXCR4+PDGFRα+ surface marker expression, and were expressing early cardiac transcription factors of both FHF and SHF progenitors.
Several studies have described co-expression of PDGFRα and low KDR (Flk1) expression, marking a tri-potent cardiac mesoderm progenitor population, which may allow monitoring of proper cardiomyocyte differentiations[46, 51, 52]. Moreover, Ardehali et al. showed that a PDGFRα+/KDR-low/ROR2+/CD13+ population could give rise to cardiomyocytes, endothelial cells, and vascular smooth muscle cells in vitro at a clonal level[53]. However, it is thought that these markers more broadly describe mesoderm and not specifically cardiac progenitors[32, 47]. Another early in vitro human CPC population that has been identified includes a tri-potent Isl1+ progenitor, which gives rise to mostly smooth muscle cells and only low numbers of cardiomyocytes[54]. These progenitors showed limited expansion in the presence of Wnt3a-producing feeders, indicating that canonical Wnt ligands can promote the expansion and maintenance of multipotent human ES- cell-derived ISL1 progenitors. Studies focussed on PSC-derived NKX2-5 expressing cardiac progenitors have shown that these were mostly bi-potent, and only rarely differentiated to endothelial cells, indicating a preceded segregation of endothelial progenitors[44, 55].

SCOPE AND OBJECTIVES OF THIS THESIS

Taken together, hPSC derived cardiac progenitor cells and cardiomyocytes offer an excellent opportunity for studying human heart development in vitro, and for translational applications, including the development of cardiovascular screening assays for drug discovery, cardiac disease modelling and safety pharmacology. However, despite high cardiomyocyte yields (up to 80-90% purity), these hPSC-derived cultures are heterogeneous cell populations. Congenital and acquired heart diseases mostly affect specific cell types, as in atrial fibrillation or dilated cardiomyopathy of the ventricles. Therefore, for disease modelling or drug discovery, it is of interest to develop protocols that specifically direct the differentiation of hPSCs towards subtypes of cardiac cells. In this thesis, we show a role for fluorescent PSC reporter lines in improving our understanding of molecular mechanisms that play a role in lineage commitment from early cardiac progenitor cells

The main objective in this thesis is: to dissect molecular mechanisms playing a role in cardiac lineage commitment, cardiac progenitor proliferation, and cardiac subtype derivation.

Chapter 2 gives a broad and detailed overview of all recent fluorescent PSC reporter lines that have been used to visualize the derivation of hPSC-derived cardiac cells, including early cardiac progenitors and subtypes of cardiomyocytes. We further explain their potential for assay development in the field of drug discovery and toxicity, and disease modelling, and we summarize their applications in the field of cellular therapy of the heart.

In **Chapter 3** we describe the generation of the dual cardiac reporter MESP1$^{mCherry/w}$NKX2-5$^{eGFP/w}$ hESC line, which allows the study of early cardiac lineage commitment. Upon cardiac differentiation, MESP1-mCherry expressing could be isolated and characterized. We identify their pre-cardiac mesodermal character and show how their further differentiation is enriched for cardiomyocytes. We further characterize how established cardiac progenitor markers are expressed in MESP1 derivatives.

Whole transcriptome analysis of MESP1-cardiac progenitor derivatives at later stages are described in **Chapter 4**. Here, we identified the sequential expression of key established cardiac transcription factors during different stages of cardiac differentiation, including early cardiac commitment (d5), early cardiac progenitor (d7), cardiac progenitor (d10), and cardiomyocyte stage (d14). We describe the identification of putative cardiac key (co)-regulators, based on their temporal expression pattern and their predicted protein-protein interactions using network-based analysis.

In **Chapter 5** we describe the generation of a dual TBX3e-mCherry-NKX2-5eGFP/w hESC reporter line. Here, we aimed to obtain nodal-like cardiomyocytes from hESCs through the use of an atrioventricular-specific enhancer and a pan-cardiac enhancer that were previously identified to synergistically activate TBX3 expression in the atrioventricular conduction system of the developing mouse embryo. A TBX3e-mCherry reporter was randomly integrated into the genome of NKX2-5eGFP/w hESCs through lentiviral delivery. We discuss both technical and biological hurdles that need to be overcome.

Chapter 6 describes the generation of a dual cardiac reporter MYL2-T2A-mCherry-NKX2-5$^{\text{eGFP/w}}$ hESC line. Here, we aimed to monitor and isolate hESC-derived ventricular cardiomyocytes, and to study the molecular mechanisms involved in their formation. We made use of the small viral T2A peptide, which, upon activation of the MYL2 promoter, results in the generation of one transcript, containing both MYL2 and mCherry coding sequences, which is translated into two separate proteins. Upon cardiac differentiation we could identify MYL2-mCherry expression cardiomyocytes, albeit at low levels. The number of MYL2 expressing ventricular cardiomyocytes could be enhanced by their dissociation and/or culture in optimized maturation medium.

In **Chapter 7**, we study the proliferative capacity of MESP1-mCherry expressing pre-cardiac progenitors. The proliferation marker cassette CAG-eGFP-Anillin was randomly integrated into the genome of MESP1$^{\text{mCherry/w}}$NKX2-5e$^{\text{GFP/w}}$ hESCs through lentiviral delivery, allowing us to study the proliferation of early cardiac progenitors up to day 5 of differentiation, prior to NKX2-5-eGFP expression. Here, we showed through time-lapse imaging how eGFP-Anillin can reflect cell division of cardiac progenitors that derive from hESCs upon directed cardiac differentiation, and we propose an algorithm-based automated image analysis set-up for small molecular and/or cytokine screenings to identify enhancing effects on cardiac progenitor and cardiomyocyte proliferation capacity.

Finally, the findings of the experimental chapters are considered as a whole in the concluding discussion in **Chapter 8**.

FIGURES

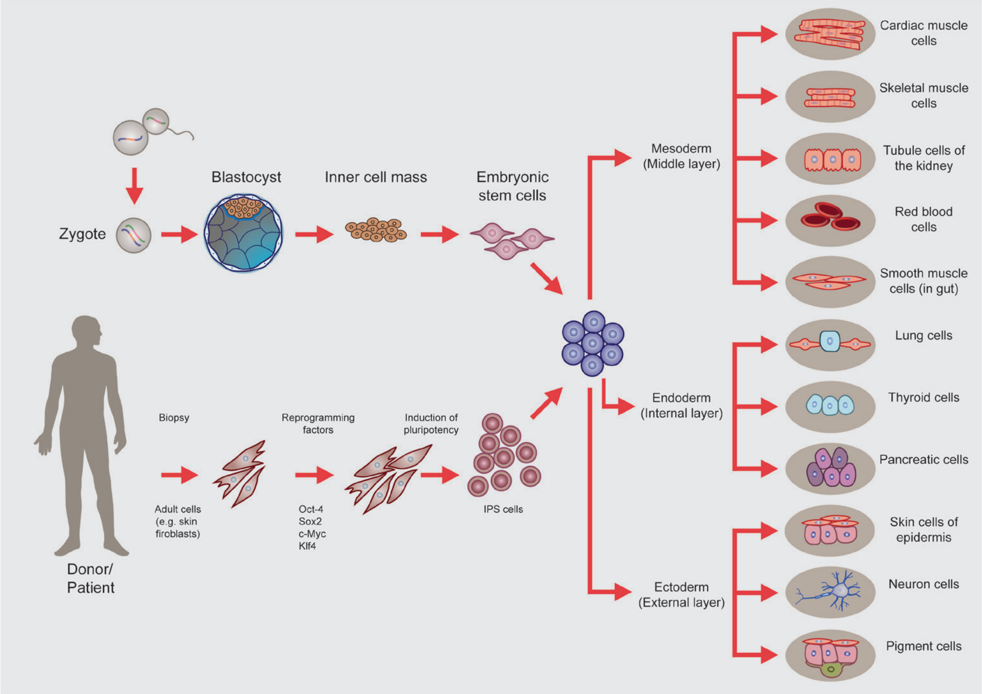


Figure 1. Human embryonic stem cells (hESCs) are derived from the inner cell mass of a blastocyst stage embryo. Human induced pluripotent stem cells (hiPSCs) are derived from human somatic cells, such as skin fibroblasts, through gene-mediated reprogramming. Both human pluripotent stem cell (hPSC) types can differentiate towards all three germ layers ectoderm, endoderm, and mesoderm, and their derivatives, including cells of the heart (cardiomyocytes).

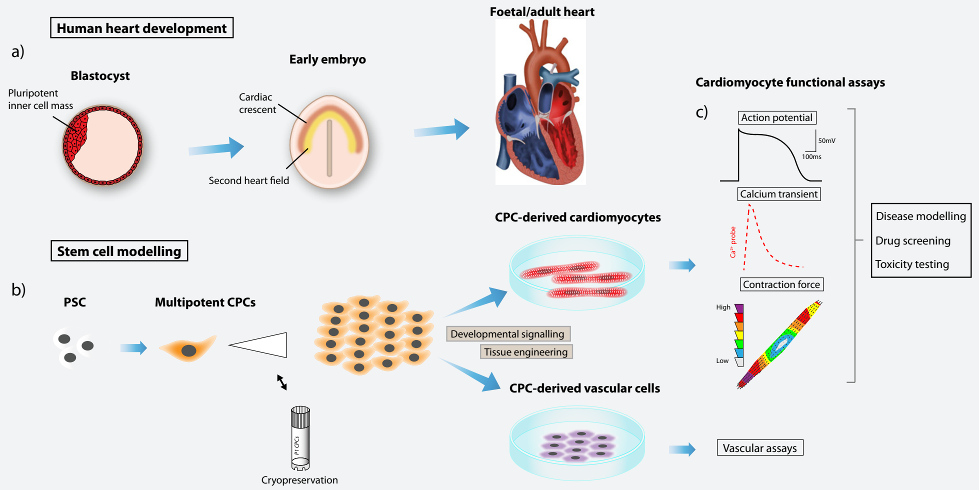


Figure 2. Pluripotent stem cell (PSC) derived cardiomyocytes can be used to study early human heart development **(A)**, but can also be applied in cardiac therapies, such as regeneration of the damaged heart **(B)**, or for experimental assay development on the field of drug discovery and toxicity screening **(C)**. *Birket et al. Developmental Biology 2015*

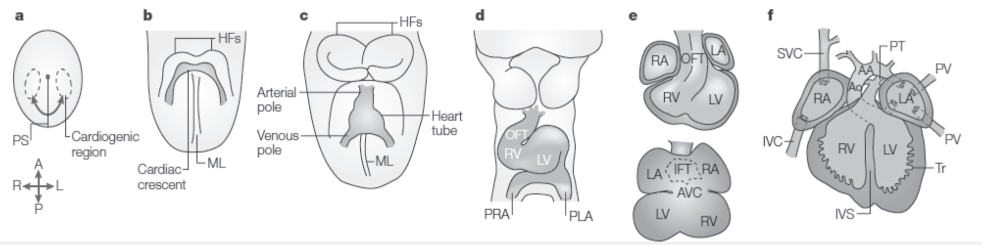


Figure 3. Morphogenesis of the mouse heart. A: Myocardial progenitor cells originate in the primitive streak (PS), from where they migrate to the anterior of the embryo at about embryonic day E6.5. **B:** These cells come to lie under the head folds (HF) and form the cardiac crescent, where differentiated myocardial cells are now observed (E7.5). **C:** The early cardiac tube forms through fusion of the cardiac crescent at the midline (ML) (E8). **D:** It subsequently undergoes looping (E8.5). **E:** By E10.5 the heart has acquired well-defined chambers, but is still a tube (upper panel, ventral view; lower panel, dorsal view). **FF** In the fetal heart (E14.5) the chambers are now separated as a result of septation and are connected to the pulmonary trunk (PT) and aorta (Ao), which ensure the separate pulmonary and systemic circulation of the blood, respectively, after birth. Deoxygenated blood enters the heart through the right atrium (RA) and is pumped to the lungs through the pulmonary trunk by the right ventricle (RV). Oxygenated blood returns to the left atrium (LA) and is pumped by the left ventricle (LV), through the aorta, to the systemic circulation that serves the whole body. Anterior (A)–posterior (P) and right (R)–left (L) axes are indicated. AA, aortic arch; AVC, atrioventricular canal; IFT, inflow tract; IVC, inferior vena cava; IVS, interventricular septum; OFT, outflow tract; PLA primitive left atrium; PRA, primitive right atrium; PV, pulmonary vein; SVC, superior vena cava; Tr, trabeculae. *Buckingham et al. Nature Reviews 2005*

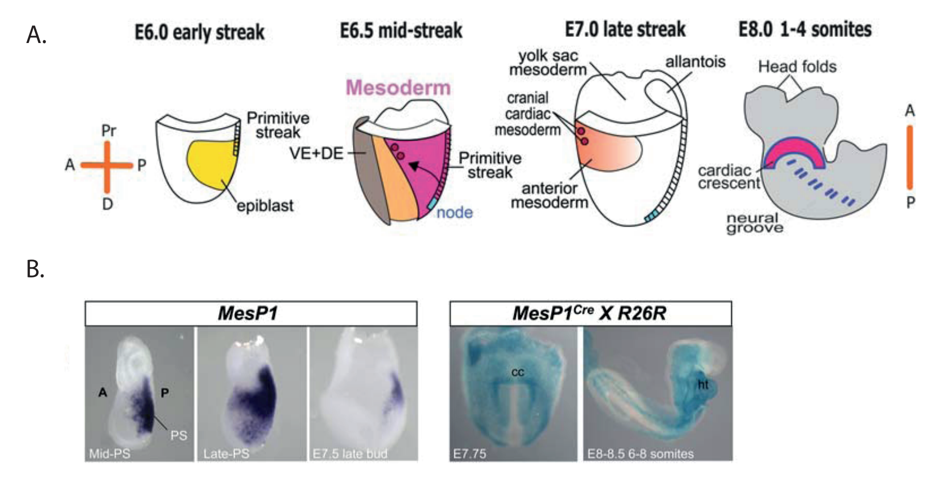


Figure 4. Early Heart Development in the Mouse. **A:** Upon early gastrulation, epiblast cells ingress through the primitive node and the primitive streak to form the ectodermal layer, and undergo epithelial-to-mesenchymal transition (EMT) to form the mesodermal and endodermal layer. The mesodermal population is marked by the expression of key transcription factor Brachyury T, from which the earliest pre-cardiac progenitor cells, expressing key transcription factor MESP1, arise from the posterior mesoderm, and migrate to the anterior-lateral regions of the embryo, at both sides of the midline. Myocardial markers are first detected at this stage. VE, visceral endoderm; DE, definitive endoderm. **B:** Time course and pattern of expression of MesP1 in early mouse embryo monitored by in situ hybridization (left panel). MesP1 Cell lineage tracing in embryos obtained from breeding MesP1-Cre with Rosa26lacZ (R26R) mice (right panel). MesP1+ cells give rise to the whole heart, as well as head and tail muscles. **C:** cardiac crescent. Ht: heart. A: anterior. P: posterior. Pr: proximal. **D:** distal. Adapted from van Vliet et al. Circulation Research 2012.

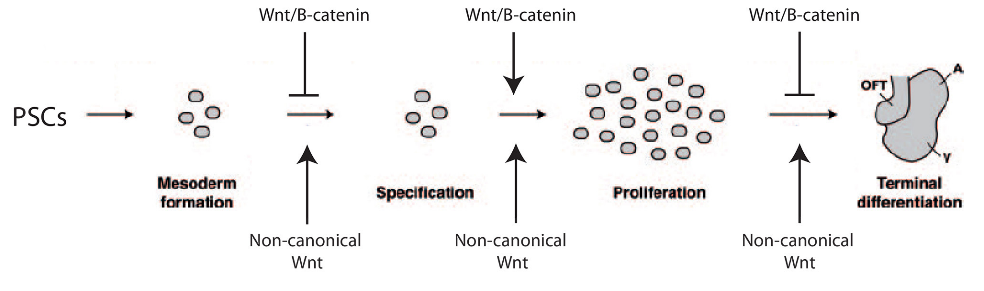


Figure 5. **The role of Wnt signalling in heart development.**

Adapted from Gessert et al. Circulation Research.

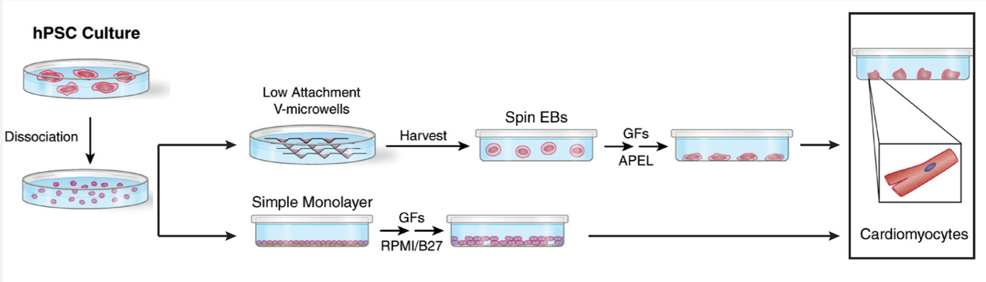


Figure 6. Principle methods of Cardiac Differentiation from human Pluripotent Stem Cells. END2 co-culture with mechanically passaged hPSCs, or spinEB or monolayer differentiation from enzymatically passaged hPSC cultures. *Adapted from Mummery et al. Circ Res 2012*

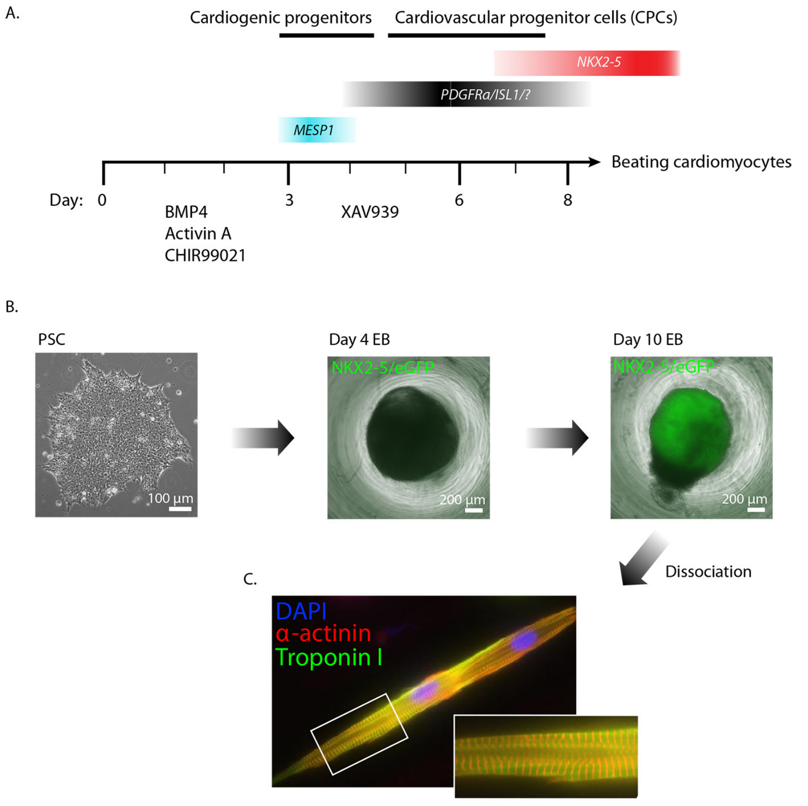


Figure 7. A: Time course of a standard monolayer differentiation protocol, based on defined factors, with the sequential expression of key cardiac genes. **B:** Example of the formation of a beating NKX2-5-eGFP expressing embryoid body (EB) at day 10 of differentiation. **C:** Staining of the sarcomeric structures Troponin I and alpha-actinin of a hPSC-derived cardiomyocyte. CHIR99021: GSK3-inhibitor; XAV939: Wnt pathway inhibitor. *Birket et al. Developmental Biology 2015*

REFERENCES

1. Thomson JA, Itskovitz-Eldor J, Shapiro SS, et al. Embryonic stem cell lines derived from human blastocysts. **SCIENCE** 1998;282(5391):1145–1147.
2. Takahashi K, Tanabe K, Ohnuki M, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. **CELL** 2007;131(5):861–872.
3. Dambrot C, van de Pas S, van Zijl L, et al. Polycistronic lentivirus induced pluripotent stem cells from skin biopsies after long term storage, blood outgrowth endothelial cells and cells from milk teeth. **DIFFERENTIATION** 2013;85(3):101–109.
4. Okita K, Yamakawa T, Matsumura Y, et al. An Efficient Nonviral Method to Generate Integration-Free Human-Induced Pluripotent Stem Cells from Cord Blood and Peripheral Blood Cells. **STEM CELLS** 2013;31(3):458–466.
5. Zhou T, Benda C, Dunzinger S, et al. Generation of human induced pluripotent stem cells from urine samples. **NATURE PROTOCOLS** 2012;7(12):2080–2089.
6. Abad M, Mosteiro L, Pantoja C, et al. Reprogramming in vivo produces teratomas and iPS cells with totipotency features. **NATURE** 2013;502(7471):340–345.
7. Chen Y, Wang K, Chandramouli GVR, et al. Trophoblast lineage cells derived from human induced pluripotent stem cells. **BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS** 2013;436(4):677–684.
8. Passier R, van Laake LW, Mummery CL. Stem-cell-based therapy and lessons from the heart. **NATURE** 2008;453(7193):322–329.
9. Finegold JA, Asaria P, Francis DP. Mortality from ischaemic heart disease by country, region, and age: statistics from World Health Organisation and United Nations. **INT. J. CARDIOL**. 2013;168(2):934–945.
10. Bergmann O, Zdunek S, Felker A, et al. Dynamics of Cell Generation and Turnover in the Human Heart. **CELL** 2015;161(7):1566–1575.
11. Chong JJH, Yang X, Don CW, et al. Human embryonic-stem-cell-derived cardiomyocytes regenerate non-human primate hearts. **NATURE** 2014;510(7504):273–277.
12. Hartogh SCD, Sluijter JP, Doevendans PA, et al. Cellular Therapy for the Infarcted Myocardium. In: Translational …. Totowa, NJ: Humana Press; 2012:341–390.
13. Braam SR, Passier R, Mummery CL. Cardiomyocytes from human pluripotent stem cells in regenerative medicine and drug discovery. **TRENDS IN PHARMACOLOGICAL SCIENCES** 2009.
14. Braam SR, Tertoolen L, van de Stolpe A, et al. Prediction of drug-induced cardiotoxicity using human embryonic stem cell-derived cardiomyocytes. **STEM CELL RESEARCH** 2010;4(2):107–116.
15. Matsa E, Burridge PW, Wu JC. Human Stem Cells for Modeling Heart Disease and for Drug Discovery. **SCIENCE TRANSLATIONAL MEDICINE** 2014;6(239):239ps6–239ps6.
16. Rajamohan D, Matsa E, Kalra S, et al. Current status of drug screening and disease modelling in human pluripotent stem cells. **BIOESSAYS** 2013;35(3):281–298.
17. Chow M, Boheler KR, Li RA. Human pluripotent stem cell-derived cardiomyocytes for heart regeneration, drug discovery and disease modeling: from the genetic, epigenetic, and tissue modeling perspectives. **STEM CELL RES THER** 2013;4(4):97.
18. Yokoo N, Baba S, Kaichi S, et al. The effects of cardioactive drugs on cardiomyocytes derived from human induced pluripotent stem cells. **BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS** 2009;387(3):482–488.

19. Mercola M, Colas A, Willems E. Induced Pluripotent Stem Cells in Cardiovascular Drug Discovery. CIRCULATION RESEARCH 2013;112(3):534–548.
20. Agarwal A, Goss JA, Cho A, et al. Microfluidic heart on a chip for higher throughput pharmacological studies. LAB CHIP 2013;13(18):3599–3608.
21. Rosenthal N, Harvey RP. Heart Development and Regeneration. Academic Press; 2010.
22. Chuva de Sousa Lopes SM, Hassink RJ, Feijen A, et al. Patterning the heart, a template for human cardiomyocyte development. DEV. DYN. 2006;235(7):1994–2002.
23. Devalla HD, Schwach V, Ford JW, et al. Atrial-like cardiomyocytes from human pluripotent stem cells are a robust preclinical model for assessing atrial-selective pharmacology. EMBO MOLECULAR MEDICINE 2015;7(4):394–410.
24. Sizarov A, Ya J, de Boer BA, et al. Formation of the Building Plan of the Human Heart Morphogenesis, Growth, and Differentiation. 2011.
25. Buckingham M, Meilhac S, Zaffran S. Building the mammalian heart from two sources of myocardial cells. NAT REV GENET 2005;6(11):826–837.
26. Larsen WJ, Sherman LS, Potter SS, et al. Human Embryology. Third. Elsevier Health Sciences; 2001.
27. Bondue A, Blanpain C. Mesp1: A Key Regulator of Cardiovascular Lineage Commitment. CIRCULATION RESEARCH 2010;107(12):1414–1427.
28. Chan SS-K, Shi X, Toyama A, et al. Mesp1 Patterns Mesoderm into Cardiac, Hematopoietic, or Skeletal Myogenic Progenitors in a Context-Dependent Manner. CELL STEM CELL 2013;12(5):587–601.
29. Yoshida T, Vivatbutsiri P, Morriss-Kay G, et al. Cell lineage in mammalian craniofacial mesenchyme. MECHANISMS OF DEVELOPMENT 2008;125(9-10):797–808.
30. Saga Y. MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube. DEVELOPMENT 1999:1–11.
31. Saga Y, Hata N, Kobayashi S, et al. MesP1: a novel basic helix-loop-helix protein expressed in the nascent mesodermal cells during mouse gastrulation. DEVELOPMENT 1996.
32. Devine WP, Wythe JD, George M, et al. Early patterning and specification of cardiac progenitors in gastrulating mesoderm. Bronner ME, ed. ELIFE SCIENCES 2014;3:e03848.
33. 33. Meilhac SM, Esner M, Kelly RG, et al. The clonal origin of myocardial cells in different regions of the embryonic mouse heart. DEVELOPMENTAL CELL 2004;6(5):685–698.
34. Lescroart F, Chabab S, Lin X, et al. Early lineage restriction in temporally distinct populations of Mesp1 progenitors during mammalian heart development. NATURE CELL BIOLOGY 2014;16(9):829–840.
35. Bondue A, Lapouge G, Paulissen C, et al. Mesp1 acts as a master regulator of multipotent cardiovascular progenitor specification. CELL STEM CELL 2008.
36. . Wu X, Golden K, Bodmer R. Heart development in Drosophila requires the segment polarity gene wingless. DEVELOPMENTAL BIOLOGY 1995;169(2):619–628.
37. Brand T. Heart development: molecular insights into cardiac specification and early morphogenesis. DEVELOPMENTAL BIOLOGY 2003;258(1):1–19.
38. Gessert S, Kühl M. The multiple phases and faces of wnt signaling during cardiac differentiation and development. CIRCULATION RESEARCH 2010;107(2):186–199.
39. Birket MJ, Mummery CL. Pluripotent stem cell derived cardiovascular progenitors--a developmental perspective. DEVELOPMENTAL BIOLOGY 2015;400(2):169–179.
40. Mummery CL, Zhang J, Ng ES, et al. Differentiation of Human Embryonic Stem Cells and Induced Pluripotent Stem Cells to Cardiomyocytes A Methods Overview. 2012.
41. Burridge PW, Matsa E, Shukla P, et al. Chemically defined generation of human cardiomyocytes. NAT METH 2014:–.
42. Lian X, Zhang J, Azarin SM, et al. Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/β-catenin signaling under fully defined conditions. NATURE PROTOCOLS 2013;8(1):162–175.
43. Mummery C, Oostwaard DW-V, Doevendans P, et al. Differentiation of human embryonic stem cells to cardiomyocytes role of coculture with visceral endoderm-like cells. CIRCULATION … 2003.
44. Elliott DA, Braam SR, Koutsis K, et al. NKX2-5eGFP/w hESCs for isolation of human cardiac progenitors and cardiomyocytes. NAT METH 2011;8(12):1037–1040.
45. Willems E, Spiering S, Davidovics H, et al. Small-Molecule Inhibitors of the Wnt Pathway Potently Promote Cardiomyocytes From Human Embryonic Stem Cell–Derived MesodermNovelty and Significance. CIRCULATION … 2011.
46. Kattman SJ, Witty AD, Gagliardi M, et al. Stage-Specific Optimization of Activin/Nodal and BMP Signaling Promotes Cardiac Differentiation of Mouse and Human Pluripotent Stem Cell Lines. CELL STEM CELL 2011;8(2):228–240.
47. Hartogh Den SC, Schreurs C, Monshouwer-Kloots JJ, et al. Dual Reporter MESP1 mCherry/w-NKX2-5 eGFP/whESCs Enable Studying Early Human Cardiac Differentiation. STEM CELLS 2014;33(1):56–67.
48. Zhang Q, Jiang J, Han P, et al. Direct differentiation of atrial and ventricular myocytes from human embryonic stem cells by alternating retinoid signals. NATURE PUBLISHING GROUP 2010;21(4):579–587.
49. Zhu W-Z, Xie Y, Moyes KW, et al. Neuregulin/ErbB Signaling Regulates Cardiac Subtype Specification in Differentiating Human Embryonic Stem CellsNovelty and Significance. CIRCULATION … 2010;(107):776–786.
50. Bondue A, Tannler S, Chiapparo G, et al. Defining the earliest step of cardiovascular progenitor specification during embryonic stem cell differentiation. THE JOURNAL OF CELL BIOLOGY 2011;192(5):751–765.
51. Kattman SJ, Huber TL, Keller GM. Multipotent Flk-1+ Cardiovascular Progenitor Cells Give Rise to the Cardiomyocyte, Endothelial, and Vascular Smooth Muscle Lineages. DEVELOPMENTAL CELL 2006;11(5):723–732.
52. Yang L, Soonpaa MH, Adler ED, et al. Human cardiovascular progenitor cells develop from a KDR+ embryonic-stem-cell-derived population. NATURE 2008;453(7194):524–528.
53. Ardehali R, Ali SR, Inlay MA, et al. Prospective isolation of human embryonic stem cell-derived cardiovascular progenitors that integrate into human fetal heart tissue. PROCEEDINGS OF THE … 2013;110(9):3405–3410.
54. Bu L, Jiang X, Martin-Puig S, et al. Human ISL1 heart progenitors generate diverse multipotent cardiovascular cell lineages. NATURE 2009;460(7251):113–117.
55. Birket MJ, Ribeiro MC, Verkerk AO, et al. Expansion and patterning of cardiovascular progenitors derived from human pluripotent stem cells. NATURE BIOTECHNOLOGY 2015;33(9):970–979.

CHAPTER 2

Fluorescent Reporters In Human Pluripotent Stem Cells: Contributions To Cardiac Differentiation And Their Applications In Cardiac Disease And Toxicity

Sabine C. Den Hartogh[1], Robert Passier[1,2]

Stem Cells. 2015 Oct 8. doi: 10.1002/stem.2196

1. Department of Anatomy and Embryology, Leiden University Medical Centre, Leiden Postal zone S-1-P, P.O. Box 9600, 2300 RC Leiden, The Netherlands
2. Department of Applied Stem cell Technologies. MIRA Institute for Biomedical Technology and Technical Medicine. University of Twente, P.O.Box 217, Enschede, The Netherlands

ABSTRACT

In the last decade, since the first report of induced pluripotent stem cells, the stem cell field has made remarkable progress in differentiation to specialized cell-types of various tissues and organs, including the heart. Cardiac lineage- and tissue-specific human pluripotent stem cell (hPSC) reporter lines have been valuable for the identification, selection and expansion of cardiac progenitor cells and their derivatives, and for our current understanding of the underlying molecular mechanisms. In order to further advance the use of hPSCs in the fields of regenerative medicine, disease modeling and preclinical drug development in cardiovascular research, it is crucial to identify functionally distinct cardiac subtypes and to study their biological signalling events and functional aspects in healthy and diseased conditions. In this review, we discuss the various strategies that have been followed to generate and study fluorescent reporter lines in human pluripotent stem cells and provide insights how these reporter lines contribute to a better understanding and improvement of cell-based therapies and pre-clinical drug and toxicity screenings in the cardiac field.

INTRODUCTION

Human pluripotent stem cells (hPSCs), i.e. embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs), can be maintained in culture indefinitely and since they have the potential to differentiate into any cell type of the human body, these cells offer an unlimited source of differentiated cell types, including cardiomyocytes (CMs). The promise of hPSC-derived CMs for treatment of heart disease has grown rapidly since the ability to reprogram somatic cells into iPSCs, discovered for the first time by Yamanaka et al. in 2006 [1]. IPSCs are very similar to embryonic stem cells (ESCs), but are of a less ethical concern, as they do not require donated embryos for derivation. Further, they can be generated from patients with genetic disease to create *in vitro* human disease models. In recent years, the differentiation efficiency from hPSCs to cardiac cells has increased considerately, yielding 60-90% cardiomyocyte cultures[2]. Efficient directed differentiation of CMs from hPSCs is of high interest for multiple fields of research, including cardiogenesis, *in vitro* disease modeling, drug discovery, safety pharmacology and regenerative medicine. Fluorescent reporter lines offer the opportunity to monitor cell subtype specification and to investigate underlying molecular mechanisms during differentiation of hPSCs towards cardiac cells, and to track cells in cell-based therapies. In order to generate fluorescent reporter lines, hPSCs need to be genetically modified, which can be performed through different approaches. A variety of fluorescent proteins that cover a broad range of the wavelength spectrum, including eGFP, mCherry and mVenus, enable imaging and tracking of specific cell populations and mapping their cell fate. Here, we will summarize the different fluorescent reporter lines that have been described so far, how these were generated, and how they have contributed to our current knowledge on cardiac progenitors segregation, proliferation, and subtype specification. Further, we describe their importance for applications in regenerative medicine (e.g. replacing damaged or non-viable heart tissue) and further advancing *in vitro* systems for disease modeling and drug discovery and safety testing **(Figure 1)**.

1 Methods for Generating PSC Reporter Lines

For selection and purification of specific cell types, non-genetic approaches such as metabolic selection or sorting using cell type-specific antibodies are preferred. However, these approaches are not always available or fail to yield sufficient purity, and therefore, the use of cell-type specific fluorescent reporter lines offers an excellent alternative. Generation of these reporter lines require genetic engineering in hPSCs, for which several successful approaches have been described. Each genetic engineering approach knows its own advantages and disadvantages, which we will discuss here briefly **(Table 1)**. More specific details can be found elsewhere[3-5].

1.1 Promoter-driven Reporter Lines

Introducing exogenous regulatory fragments that are coupled to a fluorescent reporter gene can generate cell-specific promoter reporter lines. These elements can be introduced randomly into the genome through viral-mediated transduction[6], electroporation, nucleofection, or chemical-based transfection. However, important shortcomings exist in using these methods. First, integration events are random and there is no control over copy number, site of integration, and gene expression level. Moreover, chromatin dynamics and epigenetic regulators may change upon lineage differentiation, resulting in the chance of transgene silencing disabling gene reporter expression[7-9]. Second, endogenous regulatory elements may interfere with the incorporated promoter/enhancer elements, so that the fluorescent reporter may not faithfully recapitulate endogenous activation of the target gene during differentiation. The use of strong insulator elements flanking the enhancer and its reporter gene could be a way to overcome transgene silencing[9, 10]. Moreover, endogenous elements themselves could be influenced by the integration of new regulatory elements, which may affect the differentiation process, cellular behavior and function. Several loci have been described to be ubiquitously expressed in hESCs and progeny, including sequences highly homologous to the mouse Rosa26 locus and the AAVS1 locus, which could be targeted through homologous recombination[11-13]. However, it should be noted that transgene silencing can be lineage-specific and the effect of gene expression following gene insertion in these loci have not been described for all mature cell-types. Third, these strategies are always based on restricted regulatory elements, often based on conserved elements between species. However, gene regulation is complex and can be regulated through combinatorial distal and proximal elements that control the precise expression pattern of a gene during development[14], and therefore detailed knowledge of specific regulatory elements and their effect on gene regulation is required. The best way to overcome these limitations is the development of a knock-in PSC line.

1.2 Targeted Genome Editing

With targeted genome editing, a reporter gene is inserted into a specific site at the gene of interest through homologous recombination[15]. In recent studies using cardiac reporter lines in hESCs, sequences following the translation start codon of the gene of interest were replaced in frame with the coding sequence of a fluorescent reporter gene, in the presence of adjacent to a positive selection cassette. A targeting vector containing these sequences are flanked by sequences homologous to the genomic regions at the 5' and 3' ends of the targeting site, required for homologous recombination[16, 17]. Sequential excision of the selection cassette is recommended in order to prevent interfering regulatory events at the targeted site[18]. One major disadvantages of this strategy is the generation of haploinsufficient hPSC lines, since in one allele the targeted gene of interest has been replaced with a fluorescent encoding protein. Therefore, it is important to have additional information on whether heterozygote expression of the target gene may affect differentiation and cellular functions. For example, studies in knockout mice or comparable strategies in other species may provide this information. Alternatively, an internal ribosomal entry site (IRES) or viral 2A peptide may be incorporated in the targeting construct in order to allow expression and translation of both the endogenous as well as the reporter genes[19]. However, in case of the use of IRES, it has been shown by many researchers that transcription of the second gene is less efficient than the first, which also depends on the strength of the endogenous promoter[20, 21]. Viral 2A could be a superior alternative to the widely used IRES. By using a short sequence, multiple proteins can be co-expressed from a single mRNA. However, the disadvantage is that due to ribosomal cleavage of the 2A sequence, several amino acids are added to the C terminus of the endogenous protein, and one amino acid is added to the N-terminus of the reporter protein, which may have an effect on the function or turnover of proteins[19, 22].

1.3 The Promise of Engineered Nucleases

Other site-specific genetic engineering methods that have raised interest in recent years are engineered nucleases, including zinc-finger nucleases (ZFNs)[23, 24], transcription activator-like effector nucleases (TALENS), or the most recently described clustered regularly interspaced short palindromic repeats (CRISPR) technology[4]. Important considerations to take into account when choosing a specific technology, include the size of the insertion sequence, the risk for off-target modifications[25], and the prospected efficiency of the targeting. CRISPR technology has shown high efficiency (up to 51-79% in hPSCs[26] and reliability (low off target effects) in site-specific amino acid replacement, and is considered a promising technique for disease modeling or for genetic repair. Moreover, compared to ZFNs or TALENs, it only requires the generation of a single guidance RNA (sgRNA) molecule, CAS nuclease mRNA, and a donor vector in case of an insertion, which makes it much less labor-intensive[5]. Jaenisch et al. have shown that the use of CRISPR/Cas9 can be successfully used for the efficient generation of fluorescent gene reporters in mice[27]. Here, Cas9 mRNA, sgRNA, and a donor vector containing a 3 Kb transgene cassette were injected into zygotes. From the blastocyst stage-derived mESCs, 33% were correctly targeted. However, limitations still exist and may be different in hPSCs; here, the insertion of large DNA constructs into the genome has not been described yet. For CRISPR/CAS9 technology, limitations include the variance in targeting efficiency, which depends on the size of the epitope or fluorescent tag, the amount of targeting vector to be injected, cell type, cell cycle phase[28], and the target locus, and will require model-specific optimization. More specific details are reviewed elsewhere[4, 29].

In summary, several aspects are important to take into account for generating fluorescent reporter lines, such as the amount of time and labor it takes to generate a line, the efficiency of targeting, the choice for a replacement or a fusion strategy, random gene insertions from lentiviral particles, the risk of silencing, the risk for off-target effects from nucleases, and the size of the DNA construct (a fluorescent protein and/or selection cassette). An overview of the genetic engineering methods that have been used to introduce reporter constructs in PSCs for cardiac development/differentiation and transplantation studies can be found in **Table 1**. So far, most cardiac fluorescent reporter PSCs have been generated through the use of lentiviral or homologous recombination approach, although it is expected that future studies will increasingly report generation of fluorescent reporter lines through the use of engineered nucleases[4, 28].

2.1 Fluorescent Reporters to Study Segregation of Early Cardiac Progenitors

In Vivo Cardiac Development

Since differentiation of hPSCs to the cardiac lineage recapitulate the early steps of *in vivo* heart development[30], a thorough understanding of the molecular cues and signaling pathways of this developmental process is essential. The earliest precursors for heart-forming cells are derived upon ingression of epiblast cells through the vertebrate primitive streak. During this event, called gastrulation, cells undergo epithelial-to-mesenchymal-transition (EMT) and form mesendoderm. They are marked by activation of transcription factors, including T-box transcription factor brachyury T (BryT) and homeodomain transcription factor MIXL1[31-33]. Another transcription factor, Eomesodermin (EOMES) is restricted to the anterior region of the primitive streak[34], and together with BryT they act upstream of bHLH transcription factor MESP1. In combination with low Nodal signaling, this results in the transition from a mesendodermal stage to a more defined pre-cardiac subpopulation, marked by MESP1[35-37]. Almost all cells of the final developing heart, but also hematopoietic and skeletal myogenic lineages, are derived from MESP1-expressing early progenitors[38, 39]. Recently, a more restricted cardiac progenitor cell population, derived from Mesp1, and marked by Smarcd3 was described[40]. However, information on the segregation and existence of intermediate cardiac progenitors is sparse. Based on recent clonal analysis studies in mice, early cardiovascular lineage specification is described to occur prior to, or during, Mesp1 expression[40, 41], which distinguishes the so-called first heart field (FHF, which forms the initial heart tube, and later the left ventricle (LV) and some of the right ventricle and atria) from the second heart field (SHF, marked by ISL1 expression, which forms the right ventricle (RV), atria, and outflow tract[42]. FHF progenitors are located within the so-called cardiac crescent, and following migration and fusion at the midline, a primitive heart tube is formed. SHF cells are located medially and posteriorly to the cardiac crescent and contribute to the formation of the heart at a later stage. In the developing human heart, ISL1-expressing progenitors are described to differentiate to different downstream multipotent heart progenitors that contribute to SHF-derivatives[43, 44]. In contrast, the highly conserved cardiac transcription factor NKX2-5 is first expressed in the cardiac crescent, but later as well in SHF-derived cardiac cells. Expression of NKX2-5 is maintained throughout the developing and adult heart[45].

In Vitro Cardiac Reporter Lines

A genetically modified mouse ES cell line, in which Mesp1 regulatory regions were coupled to a GFP reporter, showed the ability to isolate and study early Mesp1-GFP cells upon cardiac differentiation[46]. They identified the cell surface markers CXCR4, Flk1, and PDGFRa to be specifically expressed in Mesp1-progenitors in mice. Single cell clone analysis demonstrated the heterogeneous expression of early cardiac genes, including Isl1, indicating early segregation of Mesp1 progenitors into different downstream multi-potent cardiac progenitors. This was confirmed by an *in vivo* study in mouse, where they showed early segregation of the FHF and SHF, during MESP1 expression[41]. Recently, we characterized for the first time human pre-cardiac MESP1-progenitors differentiating to the cardiac lineage[16]. We performed targeting of the genomic locus of MESP1 through homologous recombination by replacing one allele with a DNA sequence encoding for the fluorescent protein mCherry. Microarray analysis of MESP1-mCherry positive cells largely revealed similar transcriptional characteristics to those in mice. However, a clear difference between mouse and human cardiac cell differentiations is that human MESP1 progenitors appear to have a more demarcated mesodermal character, as only a small subset of early cardiac genes is expressed in this population, including GATA4, MYL4, and MYL7. Interestingly, we also identified the absence of surface markers that were previously identified on mouse Mesp1-GFP progenitors[46]. Human MESP1 progenitors do highly express PDGFRa protein, but not CXCR4, and only a low percentage expresses KDR, indicating the presence of a PDGFRa+KDR- cardiac progenitor population, besides the previously described human KDR+PDGFR+ cardiac progenitor[47].

As described above, the LIM-homeodomain transcription factor Isl1 is enriched in the secondary heart field[42]. Mouse multipotent Isl1+ cardiac progenitors contribute to endothelial, endocardial, smooth muscle, the conduction system, right ventricular, and atrial cardiomyocytes[43, 48]. By generating an *in vitro* Isl1-Cre Knock-in DsRed reporter line in hESCs for lineage tracing construct it was found that human ISL1+ progenitors can give rise to cardiomyocytes (cTnT), endothelial cells (PECAM1), and smooth muscle cells (SM-MHC)[44]. Furthermore, this study, and several others, showed expansion of Isl1+ progenitors under activation of the Wnt/B-catenin-pathway[44, 49, 50]. More insight in the ability to maintain and expand cardiac progenitors may provide an unlimited source of cells committed to form cardiovascular cells.

As mentioned previously, NKX2-5 is a key cardiac transcription factor and is expressed in heart precursor cells in the FHF, which are committed to the cardiac lineage[51]. Elliott et al. designed a NKX2-5 fluorescent reporter line in hESCs, by targeting a DNA sequence encoding for eGFP into the NKX2-5 genomic locus through homologous recombination[17]. Since NKX2-5 is activated in cardiac progenitors and expression is maintained in functional cardiomyocytes at later stages, this reporter line is very suitable for studying the effect of small molecules and growth factors during cardiac lineage differentiation from hESCs. Indeed, following this approach, a more efficient, robust and defined differentiation protocol was attained. In addition, Birket et al. recently showed in a MYC-inducible system expansion of hESC-derived NKX2-5-eGFP+ progenitors for many passages through IGF-1 and hedgehog signaling, from which different cardiac subpopulations, including nodal-like cells, could be obtained[52]. Furthermore, using antibody arrays and gene-expression arrays, a cell surface protein signal regulatory protein (SIRPA/CD172a) was identified as a marker for early cardiac progenitor cells. Upon further differentiation, co-expression of SIRPA with VCAM1 was identified to specifically enrich for cardiomyocytes[17, 53]. Interestingly, SIRPA is not expressed in the developing mouse heart, in contrast to the human heart, which highlights the need to define markers in the human system. The previously described MESP1-mCherry targeting was performed in this NKX2-5-eGFP hESC line, yielding a dual cardiac reporter line **(Figure 2)**[16], which facilitates studying molecular pathways that are involved in the early switch from hESCs to pre-cardiac mesoderm, and its subsequent differentiation towards the cardiac lineage and specification into different cardiac subtypes. This could also provide opportunities for monitoring cardiac-committed progenitors (MESP1+) to NKX2-5+ cardiomyocytes following transplantation in infarcted hearts (see section 3.1. Fluorescent Reporters to Track Cells after Transplantation).

2.2 Fluorescent Reporters to Study Cardiomyocyte Derivation

The major cell types of the heart include ventricular and atrial cardiomyocytes, cardiac conduction cells, vascular smooth muscle cells, endothelial cells (myocardium and endocardium), and cardiac fibroblasts. Both for cardiac regeneration therapies and biomedical applications it is important to establish defined scalable culture systems or tissue constructs consisting either of pure cardiac cell populations or a well-balanced and controlled mixture of cardiac cells resembling the composition and architecture of the human heart.

However, current cardiac differentiation protocols result in a mixture of cardiac and non-cardiac cell-types. Approaches to purify cell populations using fluorescent reporter lines under the control of cardiac- and/or subtype-specific promoters have been first performed in mESCs[54-57]. In hESC an alpha-MHC or NCX-1 promoter, coupled to either a fluorescent protein (eGFP, mCherry), or the antibiotioc resistance gene, puromycin has been introduced into the genome [6, 58-60]. In the study of Kita-Matsuo et al., α-MHC positive CMs were selected in the presence of puromycin for 36 hours (day 12 to day 13.5), which yielded 92-96% pure CMs (based on quantifications of cardiac Troponin I immunostainings)[6]. Differentiation of NCX1-eGFP hESCs resulted in a maximum of ±50% of NCX-1-eGFP+ CMs, highly enriched for early and late cardiac genes, including NKX2-5, TBX5, cardiac troponin T, MYH6, and MYH7[59]. However, as alpha-MHC and NCX-1 are both expressed in all types of CMs, this results in a mixture of nodal, atrial-, and ventricular-like CMs. However, Both NCX-1 (Na+/Ca2+exchanger) and alpha-MHC are widely expressed in all subtypes, resulting in a mixture of nodal, atrial-, and ventricular-like CMs [6, 59].

2.2.1 Ventricular Cardiomyocytes

In order to select ventricular cardiomyocytes, promoter/enhancer elements of ventricular-specific (or –enriched) genes, such as ventricular myosin light chain 2v (MLC2V) or IRX4 could be appropriate candidates for generating subtype-specific reporter lines[61]. MLC2V expression is restricted to the ventricular segment of the developing human heart, followed by expressed in the entire ventricular compartment in the adult heart[62]. It is noteworthy that different cis-acting regulatory elements may be responsible for MLC2V expression in either the left or right ventricle, indicating differences in the regulation of their derivation and the possibility to identify and purify left (predominantly FHF- derived) or right ventricular (predominantly SHF-derived) cardiomyocytes[45, 63]. Huber and colleagues generated hESCs stable lines, using lentiviral vectors that expressed fluorescent protein eGFP under the transcriptional control of a highly conserved 250 bp MLC2V promoter element[64]. This enabled a highly enriched population of ventricular-like cells, purified by FACS[65]. It should be noted that the used promoter is dominantly active in right ventricular CMs[64]. Alternatively, the use of a drug resistance gene under the control of a ventricular-specific promoter may be another strategy. A disadvantage of selecting ventricular CMs through the MLC2V promoter is its low expression level in early stage hPSC-CMs. Substantial MLC2V expression levels are displayed only after long-term cultures for approximately a month, which is correlated to a higher degree of maturation [66]. Interestingly, most CM differentiation protocols that exist show a preferred differentiation to the ventricular lineage. However, it is not known whether these ventricular cells are derived from the same or from different cardiac progenitor cells, and whether it is important for development of directed differentiation protocols. The use of dual or multi-reporter hPSC lines may provide better insights in this matter.

2.2.2 Atrial Cardiomyocytes

Factors that determine fate specification of atrial cardiomyocytes are largely unknown[67], although several atrial-specific or –enriched genes have been identified previously in different species. For example, myosin light chain 2A (MLC2A) is initially expressed throughout the tubular heart in the mouse embryo, and becomes only confined to the atrial segment during chamber formation[45]. This is in contrast with human cardiac development, where MLC2A transcript was found in both atria and ventricles at even later stages of development[62], signifying the importance to understand the molecular mechanisms that are involved at the different stages of human cardiac development in order to serve as a blueprint for directed step-wise cardiac differentiation of hPSCs. Recently, the orphan nuclear receptor COUPTFII was identified as an important regulator of atrial development in mice[67]. Upon *in vitro* cardiac differentiation from hPSCs in the presence of retinoic acid, we recently demonstrated that both COUPTFI and COUPTFII are involved in atrial cardiomyocyte derivation[68]. Moreover, both transcription factors were highly enriched in human atria at RNA and protein levels during cardiac development and both genes may represent interesting candidates for generating atrial-specific reporter hPSC lines. However, expression of COUPTFII can also be found in venous/lymphatic endothelium, vascular smooth muscle cells, the endocardium, and the epicardium[69].

2.2.3 Nodal-Like Cardiomyocytes

For understanding molecular pathways leading to differentiation of nodal-like hESC-CMs and to allow their isolation, promoter-enhancer elements of several genes involved in the development of nodal tissue have been used so far. Previously, *in vivo* studies in mice showed selective activation of a chicken GATA6 (cGATA6) proximal promoter-enhancer element in the atrioventricular (AV) node and the bundle of His of the adult mouse heart[70]. Zhu and colleagues coupled this cGATA6 regulatory element to eGFP in a lentiviral construct and subsequently transduced hESCs prior to differentiation (with a transduction efficiency of only ~50%). By adjusting cardiac differentiation via inhibition of Neuregulin-1/ErbB signaling, an increased proportion of nodal-like CMs were identified[71]. cGATA6-eGFP positive cells exhibited action potential characteristics of a nodal-phenotype, indicated by the increased expression of HCN4, responsible for the funny current If, and transcription factor TBX3, important for *in vivo* specification of nodal tissue. Studies in mouse tested Hcn4 promoter/enhancer elements to investigate the derivation of nodal-like CMs, *in vitro* [72, 73]. Differentiating Hcn4-promoter-GFP mESCs resulted in only a small percentage of Hcn4 positive cells (±1%) exhibited a pacemaker-like action potential after fluorescent-based isolation[72]. Direct reprogramming of HCN4-GFP fibroblasts, using a combination of four transcription factors (Tbx3, Tbx5, Gata6, and Rxra), robustly activated HCN4p-GFP expressing cells, but without displaying electrical activity and no expression of sarcomeric proteins[73]. Alternative approaches for directed differentiation including the use of RNA molecules, such as modified mRNA, miRNAs, or lncRNAs are currently evaluated. Important to note is that at the cardiac crescent stage, Hcn4 is also expressed in early FHF progenitors and immature cardiomyocytes[73-76], and only becomes confined to components of the conduction system at later stages of development[77], which is of critical importance to realize when studying the derivation of human conduction cells.

3 The Promises and Challenges of hPSC-derived CMs for Cellular Therapy

Initially, researchers focused on the potential of hPSC-derived CMs for cell replacement therapy in experimental models of heart disease, such as myocardial infarction[78, 79]. In order to restore heart function and maintain long-term beneficial effects, sufficient transplanted cells need to engraft at the site of injury, receive blood supply, and need to couple and communicate with cardiac cells of the host myocardium in order to ensure synchronous electrophysiological and contractile function. CMs derived from human iPSCs would even provide an autologous source of cells for future cardiac regenerative therapy[80]. However, heart repair by CM replacement appeared to be challenging and many hurdles need to be taken. These promises and challenges for hPSC-derived cellular therapies for the treatment of heart failures have been reviewed elsewhere[78, 79, 81]. One crucial aspect that we would like to emphasize here, is the composition of the transplanted cell population. Which cell type(s) or tissue constructs would be most ideal regarding cell survival, tissue repair, engraftment, and functional coupling? For the delivery of organized cardiac tissue we would require scalable cultures of different cell types, such as endothelial cells, smooth muscle cells, cardiac fibroblasts, and cardiomyocytes[82]. Another important aspect is the immature phenotype of hPSC-CMs. It is expected that a higher degree of maturation of transplanted cardiomyocytes and thus a higher resemblance to the native myocardium would pose a less arrhythmogenic risk[83]. For cellular therapy, these major basic research challenges are important to be overcome.

3.1 Fluorescent Reporters to Track Cells after Transplantation

In cellular transplantation studies, it is crucial to monitor cell delivery, homing, survival, proliferation, and migration of transplanted cells[84]. Another important aspect is electrical coupling of transplanted cells with host myocardium at single cell level. Impaired communication between transplanted and host cardiomyocytes may lead to electrophysiological disturbances and consequently cardiac arrhythmias. In order to study both homing and electrical coupling, the use of a dual reporter, combining bioluminescence imaging and a fluorescent reporter under the control of a cardiac promoter has been suggested[85]. However, most studies have used a single reporter line expressing either GFP or LacZ in hPSC-derived cells[80, 83, 86, 87]. Alternatively, direct cell labelling with fluorescent probes prior to transplantation, or local or systemic injections using non-invasive radionuclide imaging, magnetic labelling, or optical mapping through bioluminescence imaging are other options[85, 88]. Limitations for cell tracking that are important to consider are the half-life of the tracking label, dilution effects during cellular division and the take-up of the label by other cells after donor cell death. More details on the advantages and disadvantages of different cell tracking techniques are reviewed elsewhere[88].

3.2 Functional Integration

Functional integration of donor cells into host myocardium can be measured by the use of fluorescent calcium dyes, such as small molecule fluo-4[86]. Calcium is a vital component of cardiac excitation-contraction coupling. A wide range of small molecule calcium dyes is available, with different excitation/emission spectra and affinities for calcium. When choosing a calcium dye, the affinity for calcium should be taken into account, as low-affinity calcium dyes will provide more accurate measurements of calcium dynamics (reviewed elsewhere)[89]. Moreover, high affinity dyes may interfere with the physiological calcium handling, affecting function of transplanted cells, and thereby not completely reflecting the final engraftment conditions. Another option is the use of genetically encoded calcium indicator proteins, which allows cell-specific calcium mapping over a long time period[90, 91]. This approach was recently followed in two studies where hESC-CMs were transplanted in infarcted non-human primate hearts or guinea-pigs[83, 92]. Fluorescent calcium indicator GCaMP3 was incorporated in hESCs at the adeno-associated virus integration site 1 (AAVS1), a region of the genome that is considered to be transcriptionally constitutive active, using zinc-finger nuclease (ZFN)-based genome editing. Here, they demonstrated electrical coupling between donor cardiomyocytes and host myocardium. However, no significant increase in ejection fraction was found and hearts suffered from ventricular arrhythmias. The main advantages and drawbacks of the different fluorescent calcium-indicators described in this section are reviewed elsewhere[89].

3.3 Cardiomyocyte Proliferation

For a long period of time, the field of cardiac regeneration has been focussing on exogenous cell delivery and endogenous activation of cardiac cells to repair the damaged heart after infarction[78, 79]. Many studies have indicated that long-term beneficial improvement of cardiac function is an immense challenge.

Although several studies indicated a low proliferation rate of adult cardiomyocytes[93, 94], it is clear that mammalian adult cardiomyocytes have lost the capacity to regenerate the heart after damage, as opposed to fish, amphibians and fetal or neonatal mammalian cardiomyocytes[95, 96], therefore it is of importance to develop strategies that may lead to proliferation of cardiac progenitors (followed by differentiation to cardiovascular cells) and/or cardiomyocytes. Analysis of cell division could be performed through the use of live mitosis markers, such as scaffolding protein anillin, which is a component of the contractile ring that localizes to specific subcellular compartments during mitosis and cytokinesis. Hesse et al. fused anillin to fluorescent protein eGFP, to visualize high spatiotemporal resolution of mitotic phase, and proved a lack of cardiomyocyte division after myocardial infarction (MI) in mouse hearts[97]. Similarly, a fluorescent anillin reporter would be very valuable for studying real-time cardiomyocyte cell division *in vitro* **(Figure 3)**. In addition, the fluorescence ubiquitination cell cycle indicator (FUCCI) is another system that enables live-imaging of cell division, based on a two-color (red and green) fluorescent proteins fused to cell cyle regulators ct1 and geminin, which are differentially expressed and ubiquinated dependent on the cell-cycle stage. This leads to a dynamic color change when cells are progressing from the G1 to S phase of the cell cycle. In contrast to an anillin reporter, the FUCCI system is not suitable for demonstrating cytokinesis[98].

4 The Promises and Challenges of hPSC-derived CMs for Disease Modeling, and Drug Discovery and Toxicity Screenings

Besides the opportunity to use human stem cell-derived cardiac cells for cell-based therapies, the utility of hPSC-CMs for disease modelling, drug discovery, and toxicity screenings has been reported in a growing number of studies, allowing researchers to study patient and/or disease-specific drug responses. hPSC-CMs are a promising alternative for animal models, as they recapitulate native human cardiomyocytes, indicated by expression of cardiac-specific genes and proteins, morphology and function. Further, they beat spontaneously, they possess calcium transients and they generate action potentials that are specific for atrial, ventricular, and nodal cells[99]. However, the use of hPSC-CMs as a platform for drug discovery need further optimization and validation and thus requires the development of tools to faithfully recapitulate cardiac physiology and function at high throughput level[100, 101]. One of the major significant challenges that scientists are facing is the mixed population of cardiac cells that are obtained from directed differentiation protocols [100, 102-104]. Therefore, it will be important to generate specific differentiation protocols in order to obtain sufficient numbers of disease-relevant subtypes of cells, which is described in detail in a previous section. Moreover, for development of advanced 3D cardiac cultures with the purpose to increase the predictability of screening assays by mimicking the heart more accurately, it is important to understand the cardiac microenvironment, including cellular and non-cellular compositions and cell-cell interactions[100].

4.1 Cardiomyocyte Electrophysiology: Voltage and Calcium Dyes

Currently published hPSC-CM disease models, including channelopathies, cardiomyopathies, and metabolic cardiac disorders, show affected (drug-induced) action and field potentials, calcium signalling, disorganized sarcomeric structures, and metabolic abnormalities such as irregular lipid or glycogen levels[105-107]. Intracellular electrophysiological recordings of single cell hPSC-CMs by patch-clamp analysis provide detailed information about their action potential and the contribution of the ion channel currents. This has gained interest, since recent development of automated medium to high-throughput whole-cell patch-clamp chip recordings[108]. An alternative approach for this labour-intensive technology is extracellular non-invasive recordings using microelectrode arrays (MEAs), which is appropriate for long-term electrophysiological analysis, and is available for high-throughput screening. This will allow analysis of CM depolarization and repolarization from multicellular field potential recordings. In addition, the use of voltage-sensitive dyes represents a technology to monitor real-time electrophysiological activity at high temporal and spatial resolution[109]. These fluorescent reporters reside within the cell membrane and are able to monitor membrane potential, based on fluctuating fluorescent intensities. As these measurements can be performed at high throughput level, voltage-sensitive dyes-based assays are of high interest for drug screening in conditions that affect electrophysiological properties, such as channelopathies or drug-induced cardiotoxicity (see below). In addition, calcium handling is another important functional parameter of CM electrophysiology, for which fluorescent dyes and reporters are available. Calcium signalling in CMs is crucial for excitation-contraction coupling. During the action potential there is a calcium influx into the cell through activation of L-type voltage-gated Ca2+ channels. This results in calcium-induced calcium release from intracellular stores of the sarcoplasmic reticulum leading to contraction. Uptake of calcium from the sarcoplasm to the sarcoplasmic reticulum (in addition to other mechanisms that lead to reduced sarcoplasmic calcium levels) ends the contraction. Fluorescent calcium dyes, also described in a section above, include small molecules such as Fluo-4, Fluo-3, or Fura-2, or the more technical challenging Fura-4F, which has the advantage of a lower affinity for calcium binding, but requires switching between emission filters when calcium concentrations change[110]. Using small molecule calcium dyes, several studies have reported affected calcium handling in hPSC-CM disease models[107]. In the majority of disease modelling studies, single parameters are measured in a sequential manner. However, in order to thoroughly understand the deeper mechanisms behind the onset of disease and its phenotype outcome, it will be of interest to measure different crucial parameters simultaneously. The combination of fluorescent reporter/dyes with non-overlapping spectra, and technological advances in LED illumination and optical filters, will make this feasible for future applications.

4.2 Cardiac Toxicity

Drug-induced cardiotoxicity of both cardiac and non-cardiac drugs is a major problem in the process of drug development and may lead to unexpected life-threatening situations[111]. For the identification of cardiotoxic effects of drugs, it is important to understand the underlying mechanisms[112]. Since current animal models and cell lines lack sufficient predictability, hPSC-CMs may represent a valuable human-based *in vitro* model for preclinical cardiotoxicity screenings. Cardiomyocyte functional readouts may include calcium and voltage levels, sarcomere contraction forces and organization, ion channel functioning, CM apoptosis, mitochondrial damage, or other putative affected signalling pathways. Cellular mechanisms that have been associated with cardiotoxicity include genotoxicity, oxidative stress, apoptosis, and lipid peroxidization[113, 114]. Automated microscopy platforms in combination with fluorescent reporters/dyes enable quantification of a variety of cardiotoxic effects, such as nuclear fragmentation or caspase 3 activity to detect apoptosis, mitochondrial membrane potentials[115-117], biochemical kinase assays based on FRET technology or specific signalling pathway analysis based on reporter assays fluorescent probes to measure cellular oxidative stress[118, 119], or fluorescent probes to measure superoxide, produced by mitrochondria under ischemic/oxidative stress conditions[120, 121]. Companies from whom such assays are available include Toxys, LifeTechnologies (Cellular Pathway Analysis Assays or CellROX), or Cyprotex. In recent years, the availability of disease-specific hPSC-CMs is growing, which allows new opportunities in the field of drug development and discovery. The demand for human-based standardized and validated cardiotoxicity assays is high and the development of reliable and sensitive pathway-specific-fluorescent reporter lines in well-characterized hPSCs in the combination with high content screening analysis methods, will be promising for medium- or high-throughput screenings of cardiotoxic side effects in the field of drug discovery[115, 122].

4.3 Optogenetics

A rapidly emerging field in biotechnology is optogenetics, which allows spatiotemporal activation of signalling cascades using green algae-derived light-sensitive cation channels, the so-called channelrhodopsins (ChR)[123]. Following activation with blue light these ChR allow the inward transfer of ions, leading to depolarization and local induction and propagation of action potentials in cardiomyocytes[124, 125]. This optogenetic approach can be followed to control electrical stimulation of heart muscle and to study subsequent cellular physiological responses under healthy and diseased conditions[123, 125, 126]. Moreover, by generating hPSC-derived channelrhodopsin-expressing nodal cardiomyocytes the realization of a light-activated bio-pacemaker may be feasible in the future. In addition, a recent innovative approach has been developed based on G-protein coupled receptors (GPCRs) containing motifs enabling light-induced spatiotemporal activation of GPCR-dependent pathways[127-129]. This will lead to a more in-depth understanding of specific intracellular signalling pathways in development, differentiation, and disease, and allows the identification of GPCRs and their downstream intracellular signalling cascades as therapeutic drugs target(s). Attributable to its ability to precisely control specific signalling pathways, optogenetics is a very promising technology for development of new assays in the field of drug development and toxicity, especially in combination with future 3D-disease modelling.

5 Cardiomyocyte Maturation and Sarcomeric Structures

One of the biggest challenges for the use of hPSC-CMs for preclinical drug screening is to increase the degree of maturity of hPSC-CMs[130]. Although morphological and functional aspects of hPSC-CMs resemble those of fetal cardiomyocytes, they still show a lack of maturity; hPSC-CMs are smaller in size, their sarcomeres, the smallest contractile units of striated muscles, have shown to be less organized when compared to adult CMs, they exhibit a lower contractile force and immature electrophysiological properties, and their metabolism is based on glycolysis, whereas adult CMs mainly use fatty acids as metabolic substrate[130, 131]. The lack of a mature phenotype may affect mechanistic and functional aspects of hPSC-CMs and consequently may result in a different drug-induced or disease-specific response. A higher degree of maturity is therefore most likely important for disease modeling and drug screening[100, 132]. Strategies to increase maturation in order to mimic human adult cardiomyocytes are followed with much interest. In addition, sarcomeric organization can be affected in hPSC-CM disease models and drug-induced cardiotoxicity[107]. It will therefore be of interest to study the organization of sarcomeric structures in hPSC-CMs with respect to the degree of maturation, and disease- and/or toxicity-induced sarcomeric disorganization **(Figure 4)**. In order to allow live-imaging for detailed analysis, a reporter line harbouring a fluorescent protein fused to sarcomeric proteins, may be a valuable additional tool, provided that the fusion protein will not affect cardiomyocyte function[133].

6 Conclusions and Future Outlook

Here, we reviewed the role of fluorescent hPSC reporter lines to our growing insights into cardiac differentiation, and their applications in the fields of cardiac regeneration, drug discovery, and toxicity screening. Fluorescent hPSC reporter lines have shown their advantages in the development and optimization of efficient cardiac differentiation protocols and will also be of high interest for deciphering the molecular signatures and mechanisms of differentiation of cardiac subtypes **(Table 2a)**, using technologies such as epigenetics (histone modifications, chromatin dynamics), transcriptomics, proteomics and metabolomics[134]. The use of dual- or multi-fluorescent reporter lines will have an additional advantage. In this manner, sequential lineage decisions can be visualized in real-time, which can be combined with in-depth molecular and functional single cell analysis[135].

A deeper understanding of the processes that determine maintenance and self-renewal of cardiac progenitors and thereby may provide an unlimited source of functional cardiac cells, will constitute an important contribution for the advancement of the fields of regenerative medicine, drug discovery and safety pharmacology. The identification of cell surface markers on specific cardiac cell subtypes allows enrichment of cardiomyocytes through cell sorting, which will be advantageous for the development of robust and validated assays for cardiomyocyte biology and/or function and for cell-based transplantation studies. In addition, improved knowledge on expansion and differentiation of cardiac progenitors may have far-reaching implications for the discovery of new therapies with a focus on activation of endogenous cardiac progenitors for the treatment of heart failure. In the context of heart repair it will also be of importance to visualize other biological processes such as cell death (necrosis and apoptosis) and the formation of cell-types following differentiation within their own lineage or as a result from switching cellular fate (e.g. via directed reprogramming). In this regard, subcellular labeling of cells using fluorescent proteins fused to specific localization signals, which target different organelles or compartments of the cell (e.g. membrane, cytosol or nucleus), offer an additional advantage for identification and tracking of transplanted cells in complex multicolor labeling experiments.

We have discussed the advantages of fluorescent reporters in cell-based transplantation studies, indicated by the possibility to visualize engrafted cells (i.e. homing and survival) and to study functional integration and maturation into the host myocardium (Table 2b). Combination of these fluorescent reporter lines with other advanced reporter lines, such as those based on optogenetics, enables studying specific signalling pathways in a temporal spatial manner. In this context the use of the light-sensitive cation channels, channelrhodopsins (ChR), derived from green algae, allow the inward transfer of ions following activation with blue light, leading to depolarization of cardiomyocyte and local induction of action potentials that are propagated to electrically coupled neighbouring cells[124, 125]. This optogenetic approach can be followed to control electrical stimulation of heart muscle and to study subsequent cellular physiological responses under healthy and diseased conditions[123, 125, 126]. Moreover, by generating hPSC-derived channelrhodopsin-expressing nodal cardiomyocytes the realization of a light-activated bio-pacemaker is one step closer. Besides the *in vivo* applications, these reporter-based strategies have an important role for the further development and optimization of human *in vitro* assays for drug discovery, safety pharmacology and disease modelling.

Finally, use of reporter-based assays and cardiac cell-types facilitates the development of three-dimensional (3D) tissues or organ-like structures for tissue engineering and generation of highly advanced models that are more closely resembling the human heart using a multidisciplinary approach. It will be intriguing to see how, among others, the emerging and widely discussed 3D bioprinting technology may contribute to achieve this goal. In addition to cardiomyocyte subtypes, additional cardiac cell-types, including vascular cells and fibroblasts and the right composition of extracellular matrix components need to be included in order to generate these structures. Clearly, many hurdles need to be overcome, but increased interest of academia and industry in the use of human stem cell-based assays **(Table 2c)** will stimulate this field and likely increases the pace of its development, which will be a big step forward in improving efficient, cost-effective and safe drug development and effective therapies for the treatment of cardiac diseases.

ACKNOWLEDGEMENTS

We thank Marcelo C. Ribeiro for acquiring CM images for Figure 4 in this review. S.C.D.H. was supported by the Rembrandt Institute for Cardiovascular Science. R.P. by ZonMw-MKMD-40-42600-98-036.

FIGURES

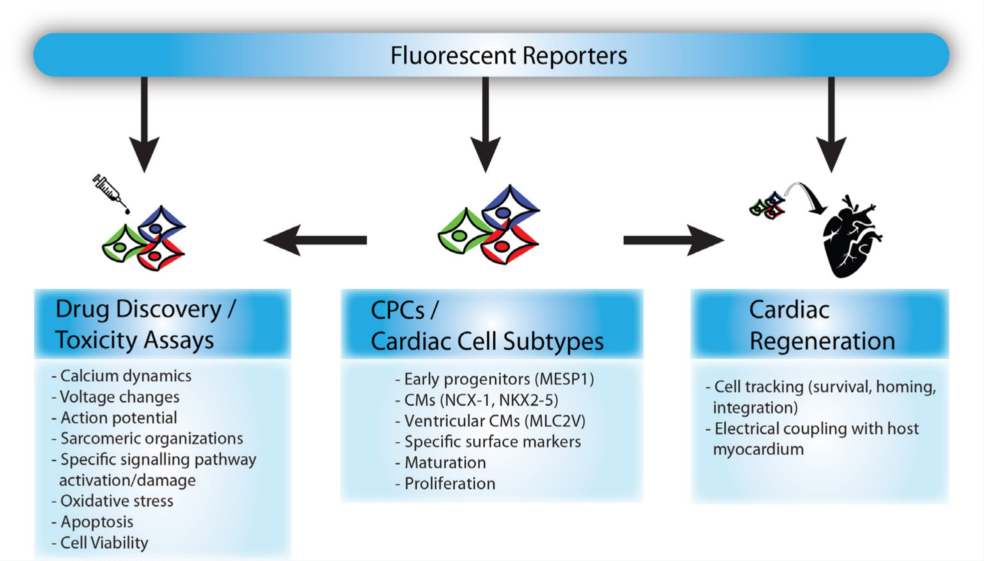


Figure 1. Schematic overview of Functional Applications of Human Fluorescent Reporter Lines for Cardiac differentiation and Cardiac Disease and Toxicity.

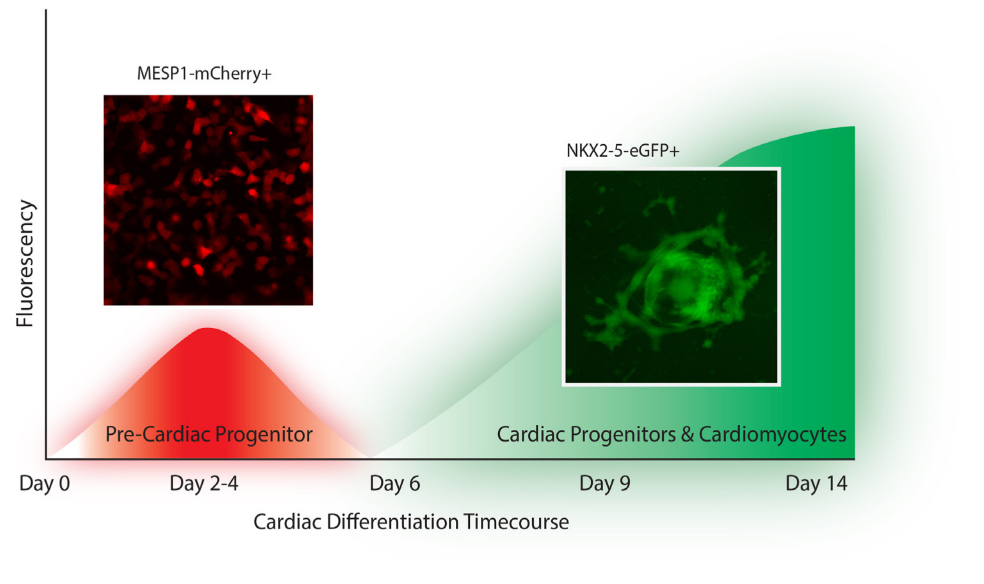


Figure 2. Dual Cardiac Fluorescent Reporter Line MESP1-mCherry-NKX2-5-eGFP in hESCs allows visualization and isolation of pre-cardiac mesoderm MESP1+ progenitor formation from hESCs, and their further differentiation to NKX2-5-eGFP+ derivatives, marking early cardiac progenitors and cardiomyocytes. *Den Hartogh et al. Stem Cells, 2015.*

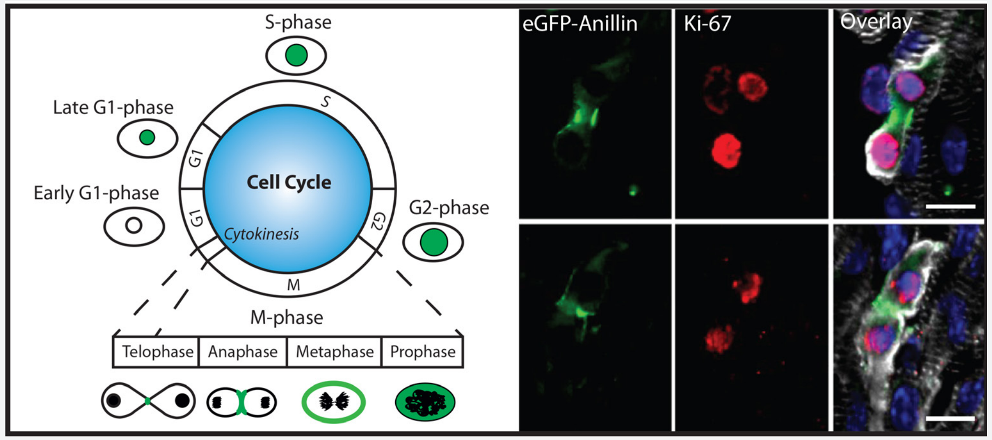


Figure 3. Fluorescent Proliferation Marker eGFP-Anillin. Left panel: Fusion of eGFP to scaffolding protein Anillin, which is differentially localized in the cell during specific cell cycle stages, allowing detection of the M-phase, and cytokinesis. Right panel: Proliferating α-actinin+ (white) cardiomyocytes in sections from an E18.5 eGFP–anillin heart are identified by eGFP labelling (green) of contractile rings or midbodies; Ki-67 staining is red; nuclei are blue. Scale bar, 10 μm. Both images were adapted from Hesse et al. Nature Communications, 2012, with permission from dr. Bernd K. Fleischmann.

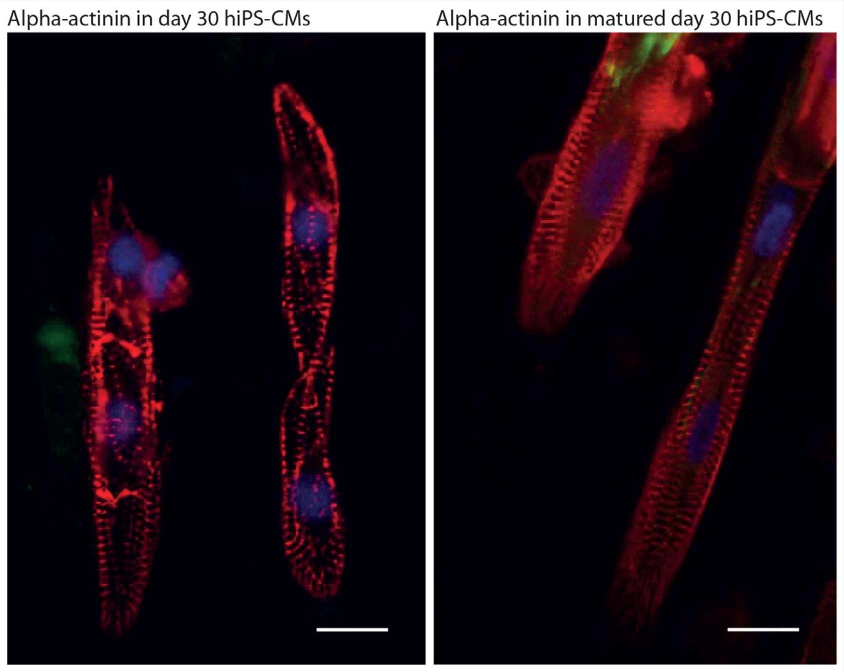


Figure 4. Sarcomeric organization in hiPS derived Cardiomyocytes improves upon maturation. Left: 30-day-old hiPSC-derived CMs cultured in standard CM medium. Right: hiPSC-derived CMs after culture in commercial media optimized for cardiomyocyte maturation. Scale bar: 15uM. Ribeiro et al., Biomaterials, 2015.

	Targeting method	**Key feature**	**Advantages**	**Disadvantages**
Knockin/Phusion	Homologous recombination with DNA donor template	- Electroporation of target construct containing large homology arms around target region	- Endogenous gene reporter expression	- Efficiency depends on locus - Disruption of one allele of the gene
	Zinc Finger Nucleases (ZFNs)	- Pair of DNA binding proteins fused to endonuclease FokI - Donor plasmid	- Endogenous gene reporter expression - Efficient and precise	- Efficiency depends on locus - Disruption of one allele of the gene
	Transcription activator-like effector nucleases (TALENs)	- Pair of DNA-binding domains fused to cleavage domain - Donor plasmid	- Endogenous gene reporter expression - Efficient and precise	- Efficiency depends on locus - Disruption of one allele of the gene
	Clustered Regulatory Interspaced Short Palindromic Repeats System (CRISPR/CAS9)	- Cas9 endonucleases - Single guidance RNA (sgRNA)	- Requires only the generation of the sgRNA - Highly efficient in site-specific amino acid replacement	- Requires optimization for large elements - Risk for off-target modifications
Random integration	BAC	- Random integration of a large genomic sequence	- Large regulatory elements can be reported - Endogenous alleles stay intact	- Instable, risk for random integration of fragmented BAC elements
	Lentivirus	- Viral packaging particles containing reporter construct	- High efficiency of integration - Rapid	- Risk for silencing, thus analysis of multiple clones required - Position effects - Random integration may disrupt other gene functions - Only for in vitro applications - No control of copy number, site of integration, and gene expression level.
	Transposons	- Mobile genetic elements	- Non-viral gene delivery - Delivery of large elements	- Risk for hyperactive transposition - Risk for integration near transcriptional start sites
	Plasmid	- Electroporation or nucleofection of plasmid	- Non-viral gene delivery	- Only small regulatory regions can be reported - Position effects

Table 1. Advantages and disadvantages of the different approaches for generating fluorescent reporter lines.

	Gene	Regulatory elements	Fluorescent/Selection tag	Cell line	Method of integration	Plasmid/virus	Reference
Mesoderm	MESP1,NKX2-5	Endogenous	*Dual:* mCherry / eGFP	NKX2-5-eGFP HES-3	HR - knockin	Targeting construct	16
	Brachyury T	Promoter fragments	eGFP	hESC	Random	Lentiviral	6
CPCs	ISL1	Endogenous	*Tracing*: cre-IRES-puro / pCAG-flox-DsRed-MST	HUES3 / H9	HR - Knockin / Random	Targeting construct / Plasmid	44
	NKX2-5	Endogenous	eGFP	HES-3 and MEL-1	HR - knockin	Targeting construct	17
CMs	MLC2V	Promoter fragment	eGFP	hESC H9.2	Random	Lentiviral	64
	α myosin heavy chain (MHC)	Promoter fragment	eGFP-IRES-Puromycin, or eGFP-IRES-herpes simplex virus thymidine kinase/-ganciclovir (HSVtk/ GCV) suicide gene system	HUES-7	Random	Lentiviral	60
	α myosin heavy chain (MHC)	Promoter fragments	eGFP, mCherry, or Puromycin	hESC	Random	Lentiviral	6
	NCX-1	Promoter fragment	eGFP	iPS (C32 and C11 /hESC (H9 and Hes3)	Random	Lentiviral	59
Nodal-phenotype	Chicken Gata6	Promoter fragment	eGFP	H7 HESCS	Random	Lentiviral	71

Table 2a. Currently published human fluorescent reporter lines in the cardiac field.

		Advantage	Disadvantage
Cell Tracking/Homing	Consititutive LacZ expression	Good for studying homing	Not informative on electrical coupling with host myocardium
	Constitutive GFP expression	Good for studying homing and integration	Not informative on electrical coupling with host myocardium
	Direct cell labeling with fluorescent probes	No need for genetic modifications	Dilution upon cell division
	Non-invasive radionuclide imaging	Sensitive	Can not be used for long-term stem cell tracking
	Magnetic labelling	Excellent soft-tissue contrast	Low sensitivity, high cost of imaging, incompatible with the use of pacemakers (if used in clinic)
	Opical mapping - bioluminiscence	Good for studying homing	
Functional Integration	Calcium dye Fluo-4	- Electrical coupling with host myocardium - Fast response time	Limited tissue penetration
	Genetically encoded GCaMP3 calcium indicator	- Cell-specific calcium mapping - Ideal calcium indicator for *in vivo* models - Electrical coupling with host myocardium	Slow response time due to slow on-off kinetics

Table 2b. Fluorescent reporter lines for cellular therapy.

Assay	Measurement
Calcium dynamics	Intracellular calcium mobilization Calcium dyes: Fluo-2/3/4/8, Fura-2, Fura-4F (ratiometric)
Transmembrane potential	Voltage sensitive dyes
Cell viability assay	Hoechst expression / DNA binding fluorescent proteins
Apoptosis assay	Caspase3/7 activity through substrate luminescence
Oxidative stress	Reactive Oxygen Species (ROS)
Cell proliferation	Ki-67 eGFP-Anillin localizations PHH3 stainings
Hypertrophy / Sarcomeric organizations	Cell size Sarcomere staining
Intracellular lipid accumulation/ Peroxidization	Quantification of fluorescent intensity in cytoplasm (perioxidization assay kits available)
Signalling pathways	In vitro fluorescent reporter assays that reveil cellular response pathways

Table 2c. Fluorescent reporter lines for disease modeling, drug discovery, and toxicity.

REFERENCES

1. Takahashi K, Yamanaka S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. **CELL** 2006;126(4):663–676.
2. Burridge PW, Keller G, Gold JD, et al. Production of de novo cardiomyocytes: human pluripotent stem cell differentiation and direct reprogramming. **CELL STEM CELL** 2012;10(1):p16–28.
3. Giudice A, Trounson A. Genetic Modification of Human Embryonic Stem Cells for Derivation of Target Cells. **CELL STEM CELL** 2008;2(5):422–433.
4. Sander JD, Joung JK. CRISPR-Cas systems for genome editing, regulation and targeting. **NAT BIOTECHNOL** 2014;32(4):347–355.
5. Harrison MM, Jenkins BV, O'Connor-Giles KM, et al. A CRISPR view of development. **GENETICS** 2014;28(17):1859–1872.
6. Kita-Matsuo H, Barcova M, Prigozhina N, et al. Lentiviral Vectors and Protocols for Creation of Stable hESC Lines for Fluorescent Tracking and Drug Resistance Selection of Cardiomyocytes. **PLOS ONE** 2009;4(4):e5046.
7. J BURKE les, BANIAHMAD A. Co-repressors 2000. **THE FASEB JOURNAL** 2000;14(13):1876–1888.
8. Xia X, Zhang Y, Zieth CR, et al. Transgenes Delivered by Lentiviral Vector are Suppressed in Human Embryonic Stem Cells in A Promoter-Dependent Manner. **STEM CELLS AND DEVELOPMENT** 2007;16(1):167–176.
9. MacArthur CC, Xue H, Van Hoof D, et al. Chromatin Insulator Elements Block Transgene Silencing in Engineered Human Embryonic Stem Cell Lines at a Defined Chromosome 13 Locus. **STEM CELLS AND DEVELOPMENT** 2012;21(2):191–205.
10. Pikaart MJ, Recillas-Targa F, Felsenfeld G. Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. **GENETICS** 1998;12(18):2852–2862.
11. Costa M, Dottori M, Ng E, et al. The hESC line Envy expresses high levels of GFP in all differentiated progeny. **NAT METH** 2005;2(4):259–260.
12. Irion S, Luche H, Gadue P, et al. Identification and targeting of the ROSA26 locus in human embryonic stem cells. **NATURE BIOTECHNOLOGY** 2007;25(12):1477–1482.
13. Hockemeyer D, Soldner F, Beard C, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. **NATURE BIOTECHNOLOGY** 2009;27(9):851–857.
14. van Weerd JH, Badi I, van den Boogaard M, et al. A large permissive regulatory domain exclusively controls tbx3 expression in the cardiac conduction system. **CIRCULATION RESEARCH** 2014;(115):432–441.
15. Davis RP, Grandela C, Sourris K, et al. Generation of Human Embryonic Stem Cell Reporter Knock-In Lines by Homologous Recombination. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2007.
16. Hartogh Den SC, Schreurs C, Monshouwer-Kloots JJ, et al. Dual Reporter MESP1 mCherry/w-NKX2-5 eGFP/whESCs Enable Studying Early Human Cardiac Differentiation. **STEM CELLS** 2014;33(1):56–67.
17. Elliott DA, Braam SR, Koutsis K, et al. NKX2-5eGFP/w hESCs for isolation of human cardiac progenitors and cardiomyocytes. NAT METH 2011;8(12):1037–1040.
18. Davis RP, Costa M, Grandela C, et al. A protocol for removal of antibiotic resistance cassettes from human embryonic stem cells genetically modified by homologous recombination or transgenesis. **NATURE PROTOCOLS** 2008;3(10):1550–1558.
19. Trichas G, Begbie J, Srinivas S. Use of the viral 2A peptide for bicistronic expression in transgenic mice. **BMC BIOLOGY** 2008;6(1):40.
20. Mizuguchi H. IRES-Dependent Second Gene Expression Is Significantly Lower Than Cap-Dependent First Gene Expression in a Bicistronic Vector. **MOL THER** 2000;1(4):376–382.
21. Hasegawa K, Cowan AB, Nakatsuji N, et al. Efficient Multicistronic Expression of a Transgene in Human Embryonic Stem Cells. **STEM CELLS** 2007;25(7):1707–1712.
22. de Felipe P, Luke GA, Hughes LE, et al. E unum pluribus: multiple proteins from a self-processing polyprotein. **TRENDS BIOTECHNOL**. 2006;24(2):68–75.
23. Urnov FD, Rebar EJ, Holmes MC, et al. Genome editing with engineered zinc finger nucleases. **NAT REV GENET** 2010;11(9):636–646.
24. Wang Y, Zhang WY, Hu S, et al. Enlightened Stem Cells in the Heart: More Efficient and Safer Reporter Gene Imaging. **CIRCULATION RESEARCH** 2012;111(12):1486–1487.
25. Veres A, Gosis BS, Ding Q, et al. Low Incidence of Off-Target Mutations in Individual CRISPR-Cas9 and TALEN Targeted Human Stem Cell Clones Detected by Whole-Genome Sequencing. **CELL STEM CELL** 2014;15(1):27–30.
26. Ding Q, Regan SN, Xia Y, et al. Enhanced Efficiency of Human Pluripotent Stem Cell Genome Editing through Replacing TALENs with CRISPRs. **CELL STEM CELL** 2013;12(4):393–394.
27. Wang H, Yang H, Shivalila CS, et al. One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. **CELL** 2013;153(4):910–918.
28. Lin S, Staahl BT, Alla RK, et al. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. Weigel D, ed. **ELIFE SCIENCES** 2015;3:e04766.
29. Ran FA, Hsu PD, Wright J, et al. Genome engineering using the CRISPR-Cas9 system. **NATURE PROTOCOLS** 2013;8(11):2281–2308.
30. Beqqali A, Kloots J, Ward-van Oostwaard D, et al. Genome-Wide Transcriptional Profiling of Human Embryonic Stem Cells Differentiating to Cardiomyocytes. **STEM CELLS** 2006;24(8):1956–1967.
31. Pearce JJH, Evans MJ. Mml, a mouse Mix-like gene expressed in the primitive streak**. MECHANISMS OF DEVELOPMENT** 1999;87(1-2):189–192.
32. Robb L, Hartley L, Begley CG, et al. Cloning, expression analysis, and chromosomal localization of murine and human homologues of aXenopus Mix gene**. DEV. DYN**. 2000;219(4):497–504.
33. Herrmann BG. Expression pattern of the Brachyury gene in whole-mount TWis/TWis mutant embryos. **DEVELOPMENT** 1991;(113):913–917.
34. Ciruna BG, Rossant J. Expression of the T-box gene Eomesodermin during early mouse development. **MECHANISMS OF DEVELOPMENT** 1999;81(1-2):199–203.
35. Costello I, Pimeisl I-M, Dräger S, et al. The T-box transcription factor Eomesodermin acts upstream of Mesp1 to specify cardiac mesoderm during mouse gastrulation. **NATURE CELL BIOLOGY** 2011;13(9):1084–1091.
36. van den Ameele J, Tiberi L, Bondue A, et al. Eomesodermin induces Mesp1 expression and cardiac differentiation from embryonic stem cells in the absence of Activin. **EMBO REPORTS** 2012:1–8.
37. David R, Jarsch VB, Schwarz F, et al. Induction of MesP1 by Brachyury(T) generates the common multipotent cardiovascular stem cell. **CARDIOVASCULAR RESEARCH** 2011;92(1):115–122.
38. Saga Y. MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube. **DEVELOPMENT** 1999:1–11.
39. Chan SS-K, Shi X, Toyama A, et al. Mesp1 Patterns Mesoderm into Cardiac, Hematopoietic, or Skeletal Myogenic Progenitors in a Context-Dependent Manner. **CELL STEM CELL** 2013;12(5):587–601.
40. Devine WP, Wythe JD, George M, et al. Early patterning and specification of cardiac progenitors in gastrulating mesoderm. Bronner ME, ed. **ELIFE SCIENCES** 2014;3:e03848.
41. Lescroart F, Chabab S, Lin X, et al. Early lineage restriction in temporally distinct populations

of Mesp1 progenitors during mammalian heart development. **NATURE CELL BIOLOGY** 2014;16(9):829–840.

42. Buckingham M, Meilhac S, Zaffran S. Building the mammalian heart from two sources of myocardial cells. **NAT REV GENET** 2005;6(11):826–837.
43. Moretti A, Caron L, Nakano A, et al. Multipotent Embryonic Isl1+ Progenitor Cells Lead to Cardiac, Smooth Muscle, and Endothelial Cell Diversification. **CELL** 2006;127(6):1151–1165.
44. Bu L, Jiang X, Martin-Puig S, et al. Human ISL1 heart progenitors generate diverse multipotent cardiovascular cell lineages. **NATURE** 2009;460(7251):113–117.
45. Rosenthal N, Harvey RP. Heart Development and Regeneration. Academic Press; 2010.
46. Bondue A, Tannler S, Chiapparo G, et al. Defining the earliest step of cardiovascular progenitor specification during embryonic stem cell differentiation. **THE JOURNAL OF CELL BIOLOGY** 2011;192(5):751–765.
47. Kattman SJ, Witty AD, Gagliardi M, et al. Stage-Specific Optimization of Activin/Nodal and BMP Signaling Promotes Cardiac Differentiation of Mouse and Human Pluripotent Stem Cell Lines. **CELL STEM CELL** 2011;8(2):228–240.
48. Laugwitz K-L, Moretti A, Lam J, et al. Postnatal isl1+ cardioblasts enter fully differentiated cardiomyocyte lineages. **NATURE CELL BIOLOGY** 2005;433(7026):647–653.
49. Qyang Y, Martin-Puig S, Chiravuri M, et al. The Renewal and Differentiation of Isl1+ Cardiovascular Progenitors Are Controlled by a Wnt/β-Catenin Pathway. **CELL STEM CELL** 2007;1(2):165–179.
50. Kwon C, Arnold J, Hsiao EC, et al. Canonical Wnt signaling is a positive regulator of mammalian cardiac progenitors. **PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES** 2007;104(26):10894–10899.
51. Sturzu AC, Wu SM. Developmental and Regenerative Biology of Multipotent Cardiovascular Progenitor Cells. **CIRCULATION RESEARCH** 2011;108(3):353–364.
52. Birket MJ, Ribeiro MC, Verkerk AO, et al. Expansion and patterning of cardiovascular progenitors derived from human pluripotent stem cells. **NATURE BIOTECHNOLOGY** 2015;33(9):970–979.
53. Dubois NC, Craft AM, Sharma P, et al. SIRPA is a specific cell-surface marker for isolating cardiomyocytes derived from human pluripotent stem cells. **NATURE BIOTECHNOLOGY** 2011;29(11):1011–1018.
54. Lee MY, Sun B, Schliffke S, et al. Derivation of functional ventricular cardiomyocytes using endogenous promoter sequence from murine embryonic stem cells. **STEM CELL RESEARCH** 2012;8(1):49–57.
55. Fijnvandraat A. Cardiomyocytes purified from differentiated embryonic stem cells exhibit characteristics of early chamber myocardium. **JOURNAL OF MOLECULAR AND CELLULAR CARDIOLOGY** 2003;35(12):1461–1472.
56. GASSANOV N, ER F, ZAGIDULLIN N, et al. Endothelin induces differentiation of ANP-EGFP expressing embryonic stem cells towards a pacemaker phenotype. **THE FASEB JOURNAL** 2004;18(14):1710–1712.
57. MULLER M. Selection of ventricular-like cardiomyocytes from ES cells in vitro. **THE FASEB JOURNAL** 2000;14(15):2540–2548.
58. Anderson D, Self T, Mellor IR, et al. Transgenic Enrichment of Cardiomyocytes From Human Embryonic Stem Cells. **MOL THER** 2007;15(11):2027–2036.
59. Ovchinnikov DA, Hidalgo A, Yang S-K, et al. Isolation of Contractile Cardiomyocytes from Human Pluripotent Stem-Cell-Derived Cardiomyogenic Cultures Using a Human NCX1-EGFPReporter. **STEM CELLS AND DEVELOPMENT** 2015;24(1):11–20.
60. Xu XQ, Zweigerdt R, Soo SY, et al. Highly enriched cardiomyocytes from human embryonic stem cells. **CYTOTHERAPY** 2008;10(4):376–389.
61. Bruneau BG, Bao Z-Z, Tanaka M, et al. Cardiac Expression of the Ventricle-Specific Homeobox Gene Irx4 Is Modulated by Nkx2-5 and dHand. **DEVELOPMENTAL BIOLOGY** 2000;217(2):266–277.
62. Chuva de Sousa Lopes SM, Hassink RJ, Feijen A, et al. Patterning the heart, a template for human cardiomyocyte development. **DEV. DYN.** 2006;235(7):1994–2002.
63. Kelly RG. Cell history determines the maintenance of transcriptional differences between left and right ventricular cardiomyocytes in the developing mouse heart. **JOURNAL OF CELL SCIENCE** 2003;116(24):5005–5013.
64. Ross RS, Navankasattusas S, Harvey RP, et al. An HF-1a/HF-1b/MEF-2 combinatorial element confers cardiac ventricular specificity and established an anterior-posterior gradient of expression. **DEVELOPMENT** 1996;(122):1799–1809.
65. Huber I, Itzhaki I, Caspi O, et al. Identification and selection of cardiomyocytes during human embryonic stem cell differentiation. **THE FASEB JOURNAL** 2007;21(10):2551–2563.
66. Lian X, hsiao C, Wilson G, et al. Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling(2). **PROCEEDINGS OF THE …** 2012;109(27):1848–1857.
67. Wu S-P, Cheng C-M, Lanz RB, et al. Atrial identity is determined by a COUP-TFII regulatory network. **DEVELOPMENTAL CELL** 2013;25(4):417–426.
68. Devalla HD, Schwach V, Ford JW, et al. Atrial-like cardiomyocytes from human pluripotent stem cells are a robust preclinical model for assessing atrial-selective pharmacology. **EMBO MOLECULAR MEDICINE** 2015;7(4):394–410.
69. Lin FJ, You LR, Yu CT, et al. Endocardial Cushion Morphogenesis and Coronary Vessel Development Require Chicken Ovalbumin Upstream Promoter-Transcription Factor II. **ARTERIOSCLEROSIS, THROMBOSIS, AND VASCULAR BIOLOGY** 2012;32(11):e135–e146.
70. Davis DL, Edwards AV, Juraszek AL, et al. A GATA-6 gene heart-region-specific enhancer provides a novel means to mark and probe a discrete component of the mouse cardiac conduction system. **MECHANISMS OF DEVELOPMENT** 2001;108(1-2):105–119.
71. Zhu W-Z, Xie Y, Moyes KW, et al. Neuregulin/ErbB Signaling Regulates Cardiac Subtype Specification in Differentiating Human Embryonic Stem CellsNovelty and Significance. **CIRCULATION …** 2010;(107):776–786.
72. MORIKAWA K, BAHRUDIN U, MIAKE J, et al. Identification, Isolation and Characterization of HCN4-Positive Pacemaking Cells Derived from Murine Embryonic Stem Cells during Cardiac Differentiation. **PACING AND CLINICAL ELECTROPHYSIOLOGY** 2010;33(3):290–303.
73. Nam YJ, Lubczyk C, Bhakta M, et al. Induction of diverse cardiac cell types by reprogramming fibroblasts with cardiac transcription factors. **DEVELOPMENT** 2014;141(22):4267–4278.
74. Vicente-Steijn R, Passier R, Wisse LJ, et al. Funny current channel HCN4 delineates the developing cardiac conduction system in chicken heart. **HEART RHYTHM** 2011;8(8):1254–1263.
75. Liang X, Wang G, Lin L, et al. HCN4 dynamically marks the first heart field and conduction system precursors. **CIRCULATION RESEARCH** 2013;113(4):399–407.
76. Späte D, Abramczuk MK, Buac K, et al. A HCN4+ cardiomyogenic progenitor derived from the first heart field and human pluripotent stem cells. **NATURE CELL BIOLOGY** 2013;15(9):1098–1106.
77. Sizarov A, Devalla HD, Anderson RH, et al. Molecular Analysis of Patterning of Conduction Tissues in the Developing Human Heart. **CIRCULATION: ARRHYTHMIA AND ELECTROPHYSIOLOGY** 2011;4(4):532–542.
78. Hartogh den S, Sluijter J, Van Laake LW, et al. Chapter 11. Cellular therapy for the infarcted myocardium. In: Translational Cardiology. Springer; 2012.
79. Passier R, van Laake LW, Mummery CL. Stem-

cell-based therapy and lessons from the heart. **NATURE** 2008;453(7193):322–329.

80. Nelson TJ, Martinez-Fernandez A, Yamada S, et al. Repair of acute myocardial infarction by human stemness factors induced pluripotent stem cells. **CIRCULATION ...** 2009;120(5):408–416.
81. Couzin-Frankel J. The elusive heart fix. **SCIENCE** 2014;345(6194):252–257.
82. Spater D, Hansson EM, Zangi L, et al. How to make a cardiomyocyte. **DEVELOPMENT** 2014;141(23):4418–4431.
83. Chong JJH, Yang X, Don CW, et al. Human embryonic-stem-cell-derived cardiomyocytes regenerate non-human primate hearts. **NATURE** 2014;510(7504):273–277.
84. Cao F, Lin S, Xie X, et al. In vivo visualization of embryonic stem cell survival, proliferation, and migration after cardiac delivery. **CIRCULATION ...** 2006;113(7):1005–1014.
85. de Almeida PE, van Rappard JRM, Wu JC. In vivo bioluminescence for tracking cell fate and function. **AM. J. PHYSIOL. HEART CIRC. PHYSIOL.** 2011;301(3):H663–71.
86. Ardehali R, Ali SR, Inlay MA, et al. Prospective isolation of human embryonic stem cell-derived cardiovascular progenitors that integrate into human fetal heart tissue. **PROCEEDINGS OF THE ...** 2013;110(9):3405–3410.
87. Van Laake LW, Qian L, Cheng P, et al. Reporter-Based Isolation of Induced Pluripotent Stem Cell- and Embryonic Stem Cell-Derived Cardiac Progenitors Reveals Limited Gene Expression Variance. **CIRCULATION RESEARCH** 2010;107(3):340–347.
88. Fu Y, Kraitchman DL. Stem cell labeling for noninvasive delivery and tracking in cardiovascular regenerative therapy. **EXPERT REV CARDIOVASC THER** 2010;8(8):1149–1160.
89. Herron TJ, Lee P, Jalife J. Optical Imaging of Voltage and Calcium in Cardiac Cells & Tissues. **CIRCULATION RESEARCH** 2012;110(4):609–623.
90. Tallini YN, Ohkura M, Choi BR, et al. Imaging cellular signals in the heart in vivo: Cardiac expression of the high-signal Ca2+ indicator GCaMP2. **PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES** 2006;103(12):4753–4758.
91. Tian L, Hires SA, Mao T, et al. Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. **NAT METH** 2009;6(12):875–881.
92. Shiba Y, Fernandes S, Zhu W-Z, et al. Human ES-cell-derived cardiomyocytes electrically couple and suppress arrhythmias in injured hearts. **NATURE** 2012;489(7415):322–325.
93. Bergmann O, Zdunek S, Felker A, et al. Dynamics of Cell Generation and Turnover in the Human Heart. **CELL** 2015;161(7):1566–1575.
94. Bergmann O, Bhardwaj RD, Bernard S, et al. Evidence for Cardiomyocyte Renewal in Humans. **SCIENCE** 2009;324(5923):98–102.
95. Takeuchi T. Regulation of cardiomyocyte proliferation during development and regeneration. **DEV GROWTH DIFFER** 2014;56(5):402–409.
96. Ikenishi A, Okayama H, Iwamoto N, et al. Cell cycle regulation in mouse heart during embryonic and postnatal stages. **DEV GROWTH DIFFER** 2012;54(8):731–738.
97. Hesse M, Raulf A, Pilz G-A, et al. Direct visualization of cell division using high-resolution imaging of M-phase of the cell cycle. **NAT COMMS** 2012;3:1076.
98. Zielke N, Korzelius J, van Straaten M, et al. Fly-FUCCI: A Versatile Tool for Studying Cell Proliferation in Complex Tissues. **CELL REPORTS** 2014;7(2):588–598.
99. Chow M, Boheler KR, Li RA. Human pluripotent stem cell-derived cardiomyocytes for heart regeneration, drug discovery and disease modeling: from the genetic, epigenetic, and tissue modeling perspectives. **STEM CELL RES THER** 2013;4(4):97.
100. Mercola M, Colas A, Willems E. Induced Pluripotent Stem Cells in Cardiovascular Drug Discovery. **CIRCULATION RESEARCH** 2013;112(3):534–548.
101. Agarwal A, Goss JA, Cho A, et al. Microfluidic heart on a chip for higher throughput pharmacological studies. **LAB CHIP** 2013;13(18):3599–3608.
102. Orlova VV, van den Hil FE, Petrus-Reurer S, et al. Generation, expansion and functional analysis of endothelial cells and pericytes derived from human pluripotent stem cells. **NATURE PROTOCOLS** 2014;9(6):1514–1531.
103. Cheung C, Bernardo AS, Trotter MWB, et al. Generation of human vascular smooth muscle subtypes provides insight into embryological origin–dependent disease susceptibility. **NATURE BIOTECHNOLOGY** 2012;30(2):165–175.
104. Burridge PW, Matsa E, Shukla P, et al. Chemically defined generation of human cardiomyocytes. **NAT METH** 2014:–.
105. Benam KH, Dauth S, Hassell B, et al. Engineered in vitro disease models. **ANNU REV PATHOL** 2015;10:195–262.
106. Wang G, McCain ML, Yang L, et al. Modeling the mitochondrial cardiomyopathy of Barth syndrome with induced pluripotent stem cell and heart-on-chip technologies. **NAT. MED.** 2014;20(6):616–623.
107. Sharma A, Wu JC, Wu SM. Induced pluripotent stem cell-derived cardiomyocytes for cardiovascular disease modeling and drug screening. **STEM CELL RES THER** 2013;4(6):150.
108. Friedrichs S, Malan D, Voss Y, et al. Scalable Electrophysiological Investigation of iPS Cell-Derived Cardiomyocytes Obtained by a Lentiviral Purification Strategy. **JOURNAL OF CLINICAL MEDICINE** 2015, VOL. 4, PAGES 102-123 2015;4(1):102–123.
109. Warren M, Spitzer KW, Steadman BW, et al. High-precision recording of the action potential in isolated cardiomyocytes using the near-infrared fluorescent dye di-4-ANBDQBS. **AM. J. PHYSIOL. HEART CIRC. PHYSIOL.** 2010;299(4):H1271–81.
110. Wokosin DL, Loughrey CM, Smith GL. Characterization of a range of fura dyes with two-photon excitation. **BIOPHYS. J.** 2004;86(3):1726–1738.
111. Patterson C, Willis MS. Translational Cardiology. Springer; 2012.
112. Force T, Kolaja KL. Cardiotoxicity of kinase inhibitors: the prediction and translation of preclinical models to clinical outcomes. **NAT REV DRUG DISCOV** 2011;10(2):111–126.
113. Hendriks G, Atallah M, Morolli B, et al. The ToxTracker Assay: Novel GFP Reporter Systems that Provide Mechanistic Insight into the Genotoxic Properties of Chemicals. **TOXICOLOGICAL SCIENCES** 2011;125(1):285–298.
114. Albini A, Pennesi G, Donatelli F, et al. Cardiotoxicity of anticancer drugs: the need for cardio-oncology and cardio-oncological prevention. **J. NATL. CANCER INST.** 2010;102(1):14–25.
115. Mioulane M, Foldes G, Ali NN, et al. Development of high content imaging methods for cell death detection in human pluripotent stem cell-derived cardiomyocytes. **J. OF CARDIOVASC. TRANS. RES.** 2012;5(5):593–604.
116. Mitra K, Lippincott-Schwartz J. Analysis of mitochondrial dynamics and functions using imaging approaches. **CURR PROTOC CELL BIOL 2**010;Chapter 4:Unit 4.25.1–21.
117. Kerkela R, Woulfe KC, Durand J-B, et al. Sunitinib-Induced Cardiotoxicity Is Mediated by Off-Target Inhibition of AMP-Activated Protein Kinase. **CLINICAL AND TRANSLATIONAL SCIENCE** 2009;2(1):15–25.
118. Hendriks G, Atallah M, Raamsman M, et al. Sensitive DsRed fluorescence-based reporter cell systems for genotoxicity and oxidative stress assessment. **MUTAT. RES.** 2011;709-710:49–59.
119. van der Linden SC, Bergh von ARM, van Vugt-Lussenburg BMA, et al. Development of a panel of high-throughput reporter-gene assays to detect genotoxicity and oxidative stress. **MUTAT RES GENET TOXICOL ENVIRON MUTAGEN** 2014;760:23–32.
120. Becker LB, vanden Hoek TL, Shao ZH, et al. Generation of superoxide in cardiomyocytes

during ischemia before reperfusion. **AM. J. PHYSIOL.** 1999;277(6 Pt 2):H2240–6.

121. Dröse S, Brandt U. Molecular mechanisms of superoxide production by the mitochondrial respiratory chain. **ADV. EXP. MED. BIOL.** 2012;748:145–169.
122. Sinnecker D, Laugwitz K-L, Moretti A. Induced pluripotent stem cell-derived cardiomyocytes for drug development and toxicity testing. **PHARMACOL. THER**. 2014;143(2):246–252.
123. Ambrosi CM, Klimas A, Yu J, et al. Cardiac applications of optogenetics. **PROGRESS IN BIOPHYSICS AND MOLECULAR BIOLOGY** 2014;115(2-3):294–304.
124. Wikswo JP, Lin SF, Abbas RA. Virtual electrodes in cardiac tissue: a common mechanism for anodal and cathodal stimulation. **BIOPHYS. J.** 1995;69(6):2195–2210.
125. Bruegmann T, Malan D, Hesse M, et al. Optogenetic control of heart muscle in vitro and in vivo. **NAT METH** 2010;7(11):897–900.
126. Boyle PM, Williams JC, Ambrosi CM, et al. A comprehensive multiscale framework for simulating optogenetics in the heart. **NAT COMMS** 2013;4:2370.
127. Makowka P, Bruegmann T, Beiert T, et al. Optogenetic GS Activation in Cardiomyocytes Enhances Pacemaker Activity. **BIOPHYS. J.** 2014;106(2):776a–.
128. Janovjak H, Szobota S, Wyart C, et al. A light-gated, potassium-selective glutamate receptor for the optical inhibition of neuronal firing. **NAT NEUROSCI** 2010;13(8):1027–1032.
129. Levitz J, Pantoja C, Gaub B, et al. Optical control of metabotropic glutamate receptors. **NAT NEUROSCI** 2013;16(4):507–516.
130. Yang X, Pabon L, Murry CE. Engineering adolescence: maturation of human pluripotent stem cell-derived cardiomyocytes. **CIRCULATION RESEARCH** 2014;114(3):511–523.
131. Ribeiro MC, Tertoolen LG, Guadix JA, et al. Functional maturation of human pluripotent stem cell derived cardiomyocytes in vitro--correlation between contraction force and electrophysiology. **BIOMATERIALS** 2015;51:138–150.
132. Kim C, Wong J, Wen J, et al. Studying arrhythmogenic right ventricular dysplasia with patient-specific iPSCs. **NATURE** 2013;494(7435):105–110.
133. Dabiri GA, Turnacioglu KK, Sanger JM, et al. Myofibrillogenesis visualized in living embryonic cardiomyocytes. **PROC. NATL. ACAD. SCI. U.S.A.** 1997;94(17):9493–9498.
134. Wamstad JA, Alexander JM, Truty RM, et al. Dynamic and Coordinated Epigenetic Regulation of Developmental Transitions in the Cardiac Lineage. **CELL** 2012;151(1):206–220.
135. Hoppe PS, Coutu DL, Schroeder T. Single-cell technologies sharpen up mammalian stem cell research. **NATURE CELL BIOLOGY** 2014;16(10):919–927.

CHAPTER 3

Dual reporter MESP1$^{mCherry/w}$-NKX2-5e$^{GFP/w}$ hESCs enable studying early human cardiac differentiation

Sabine C. Den Hartogh[1], Chantal Schreurs[1], Jantine J. Monshouwer-Kloots[1], Richard P. Davis[1], David A. Elliott[2], Christine Mummery[1], Robert Passier[1].

Stem Cells. 2015 Oct 8. doi: 10.1002/stem.2196

1. Department of Anatomy and Embryology, Leiden University Medical Centre, Leiden Postal zone S-1-P, P.O. Box 9600, 2300 RC Leiden, The Netherlands
2. Murdoch Childrens Research Institute, Royal Children's Hospital, Parkville, Victoria, Australia

ABSTRACT

Understanding early differentiation events leading to cardiogenesis is crucial for controlling fate of human pluripotent stem cells and developing protocols that yield sufficient cell numbers for use in regenerative medicine and drug screening. Here, we develop a new tool to visualize patterning of early cardiac mesoderm and cardiomyocyte development *in vitro* by generating a dual MESP1mCherry$^{/w}$-NKX2-5$^{eGFP/w}$ reporter line in human embryonic stem cells (hESCs) and using it to examine signals that lead to formation of cardiac progenitors and subsequent differentiation. MESP1 is a pivotal transcription factor for pre-cardiac mesoderm in the embryo, from which the majority of cardiovascular cells arise. Transcription factor NKX2-5 is expressed upon cardiac crescent formation. Induction of cardiac differentiation in this reporter line resulted in transient expression of MESP1-mCherry, followed by continuous expression of NKX2-5-eGFP. MESP1-mCherry cells showed increased expression of mesodermal and EMT (epithelial-mesenchymal-transition) markers confirming their mesodermal identity. Whole-genome microarray profiling and FACS analysis of MESP1-mCherry cells showed enrichment for mesodermal progenitor cell surface markers PDGFR-α, CD13, and ROR2. No enrichment was found for the previously described KDR+PDGFR-α+ progenitors. MESP1-mCherry derivatives contained an enriched percentage of NKX2-5-eGFP and Troponin T expressing cells, indicating preferential cardiac differentiation; this was enhanced by inhibition of the Wnt-pathway. Furthermore, MESP1-mCherry derivatives harboured smooth muscle cells and endothelial cells, demonstrating their cardiac and vascular differentiation potential under appropriate conditions. The MESP1-NKX2-5 hESC reporter line allows us to identify molecular cues crucial for specification and expansion of human cardiac mesoderm and early progenitors and their differentiation to specific cardiovascular derivatives.

INTRODUCTION

Human pluripotent stem cells (hPSCs) offer exceptional opportunities for regenerative medicine, drug discovery and toxicity, disease modelling and developmental biology because of their ability to differentiate in principle into any cell type of the human body. In particular, their differentiation towards cell types that are challenging to obtain and maintain as primary cultures from tissues with limited or no self-repair capacity, such as the brain and heart, have made them attractive as research models. Previously, we and others have shown that cardiomyogenesis of hPSCs *in vitro* recapitulates signalling events and sequential transcriptional activation that occurs during heart development *in vivo* [1, 2]. Moreover, as knowledge of the regulatory mechanisms that govern *in vivo* cardiac development has advanced, it has been possible to translate this to differentiation protocols based on stepwise temporal modification of signalling pathways. This has led to a remarkable increase in the yield of cardiac cells not only from hESCs but also from human induced pluripotent stem cells (reviewed in [3]).

However, despite these advances, the cardiac derivatives obtained contain heterogeneous populations of cardiac cells: not only different subtypes of cardiomyocytes but also endothelial cells and smooth muscle cells. This is because the molecular cues underlying specification of early human cardiac progenitors towards specific cardiac subtypes have not been fully elucidated. For future clinical and pharmaceutical applications, better insight into the differentiation of cardiac cell subtypes is necessary.

Several studies have shown that different cardiac subtypes are derived from multiple early cardiac progenitors that segregate soon after gastrulation[4-6]. The molecular mechanisms that play a role in the formation of early cardiac progenitors however, remain largely unknown[4, 7]. In order to study the onset of cardiac differentiation in human, it is necessary to identify early populations from common progenitors at the mesodermal stage of development and identify the molecular pathways involved in the further differentiation of the cardiac lineage.

Mesoderm transcription factor MESP1 is important in this context as *in vivo* lineage tracing in mouse has demonstrated that almost all cells of the heart are derived from a common MESP1-expressing mesodermal progenitor, including myo-, endo, and epicardium, and cells of the conduction system[8-10]. MESP1 is transiently expressed in the anterior mesoderm from the onset of gastrulation prior to cardiac crescent formation[8, 11, 12]. Furthermore, MESP1 overexpression in mouse ESCs showed that MESP1 induces the expression of cardiovascular genes[13-15]. A recent study in mouse illustrated that besides the cardiac lineage, MESP1 derivatives also encompass cells of the hematopoietic- and skeletal myogenic lineages, which emerge depending on the endogenous and exogenous cues present during MESP1 expression[16]. Of note, all these studies have been performed in mice. However, to instigate advances in human heart development, regenerative medicine, disease modelling and drug testing, it is necessary to understand the earliest steps of cardiac differentiation in a human model.

In order to identify, isolate and characterise MESP1 expressing cells during human cardiomyogenesis, we introduced the mCherry reporter into the MESP1 locus of the NKX2-5eGFP hESC reporter line generated previously [17], leading to a dual reporter system for these early cardiovascular lineage markers in hESCs. We were able to isolate the MESP1-expressing mesoderm progenitors in a cardiac-inducing environment and demonstrate their ability to differentiate further to a highly enriched NKX2-5eGFP expressing cardiac population. We further confirmed the presence of other mesoderm and cardiac progenitor populations in MESP1-mCherry expressing-derivatives, including CD13+ROR-2+ and PDGFR-α+KDR+ populations [18-20]. Although CD13, ROR-2, and PDGFR-α expression were strongly enriched, they were not specifically expressed in the MESP1-mCherry derivatives. In addition, no enrichment was found for the previously described PDGFR-α+KDR+ population. Finally, we demonstrate that this MESP1^{mCherry}-NKX2-5^{eGFP} dual reporter hESC line is a useful tool to study the differentiation of pre-cardiac mesoderm towards early cardiac progenitor cells and their derivatives. It now provides an opportunity to study the molecular mechanisms underlying the onset of progenitor subtypes formation specifically in the cardiac lineage in human. This hESC dual reporter will contribute to providing better insights into cardiomyocyte determination and form the basis for improved differentiation protocols for specifying cardiac subtypes (atrial, ventricular, pacemaker cells). This will be relevant to future pharmaceutical and clinical applications of specific cardiac cell subtypes.

METHODS

MESP1 Gene Targeting in hESCs

To generate the *MESP1* targeting vector, the bacterial artificial chromosome (BAC) RP11-975L13 (Source Bioscience, Lifesciences) was modified by recombineering (Gene Bridges). The mCherry cassette and a selection marker replaced the MESP1 exon 1 sequence (oligonucleotides used for recombineering: MESP1_5arm_PvuI and MESP1_3_arm, **Supp. Table S1**). The modified *MESP1* locus and the surrounding 5'and 3'homology arms were subsequently subcloned into the *MESP1* targeting vector. The final vector comprised a 8.6-kilobase (kb) 5'homology arm, mCherry, *loxP*-flanked G418 resistance cassette and a 2.3-kb 3'homology arm.

The targeting vector was linearized by PvuI digestion and 20 µg was electroporated into 9x106 NKX2-5-eGFP reporter hESCs (HES3-NKX2-5$^{eGFP/w}$)[17] as described previously[21]. Targeted clones were identified by PCR using oligonucleotides specific for mCherry (P1) and a region immediately 3'of the targeting vector (P2) (**Fig. S1, Table S1**).

Targeted clones were clonally derived and transfected with Cre-recombinase to excise the *loxP*-flanked neomycin-resistance cassette as described previously[22]. Neomycin excision was confirmed by PCR (**Fig. S1, Table S1**). Undifferentiated, targeted clones were fixed in 2% paraformaldehyde and were used for immunocytochemistry (Nanog, Tra-181, and Oct-4) to confirm their pluripotent status. For karyotyping, 20 metaphase chromosome spreads were examined for each line.

Single Cell Sorting

For single cell sorting, hESCs were enzymatically passaged using 1x TrypLE Select (Life Technologies). After dissociation, cells were resuspended in FACS sorting buffer (PBS with 0.5% BSA and 2.5mM EDTA) and filtered through a 40-µm cell strainer (Falcon). Cells were sorted using a BD ARIA III flow cytometer into 96 wells flat-bottom low attachment plates, live cells were gated on the basis of side scatter, forward scatter and propidium iodide exclusion. Flow cytometric gates were set using control cells labelled with the appropriate isotype control antibody.

hESC Culture and Cardiac Differentiation

MESP1$^{mCherry/w}$ *NKX2-5*$^{eGFP/w}$ hESCs were cultured on mouse embryonic fibroblasts on 2.85 cm2 organ dishes (Falcon), in 1 mL hESC medium (DMEM F12, Non Essential Amino Acids, Knock Out Serum (Gibco by Life Technologies); bFGF (Miltenyi Biotec)) and were mechanically passaged[23]. One or two days before differentiation, cells were passaged onto a 12-well tissue culture dish coated with Matrigel (BD Biosciences; growth factor reduced, phenol red–free), in 1 mL of hESC medium per well. On the first day of differentiation, day 0, hESC medium was replaced by the low insulin serum-free medium BPEL (BSA, polyvinyl alcohol, essential lipids, as previously described)[17, 24], containing the cytokines BMP4 (30 ng ml^{-1}, R&D Systems) and Activin A (30 ng ml^{-1}, Miltenyi Biotec), and Chir99021 (1.5uM, Axon Medchem). At day 3 of differentiation, the medium was refreshed with BPEL containing 5uM Xav939 (R&D). On day 7 of differentiation, the medium was replaced with BPEL without growth factors.

Isolation and Re-aggregation of MESP1-mCherry cells

On day 3 of differentiation, cells were sorted by flow cytometry as described above, based on mCherry expression (both positive and negative populations). Flow cytometric gates were set using undifferentiated hESCs. Sorted cells were seeded in a 96-well V-shaped non-adherent plate (Greiner Bio One) at 20,000 cells per well in 50 µL BPEL in the presence of Xav939 (5uM, R&D) to enhance cardiac differentiation. Aggregation was initiated by a brief centrifugation step for 3 minutes at 480g. Four days after replating, aggregates were refreshed with BPEL without addition of growth factors or inhibitors.

Analysis of MESP1-mCherry Derivatives

Cardiomyocytes derived from MESP1-mCherry-sorted cells were analyzed by flow cytometry and immunohistochemistry (IHC) on day 3, 5 and day 14 of differentiation (0, 2 and 11 days after replating). For IHC, MESP1-mCherry-derived aggregates were fixed in 2% PFA at day 14. For alpha-actinin and SM-HC IHC, MESP1-mCherry-derived aggregates were dissociated using 1x TrypLE Select, plated on matrigel in BPEL, and fixed after four days.

For flow cytometric analysis, aggregates were collected, dissociated as described before, and resuspended in FACS washing buffer (PBS with 2% FCS and 2.5mM EDTA). For analysis of intracellular markers, cells were treated with FIX & PERM® Cell Fixation and Permeabilization Kit (Life Technologies). Analyses were carried out using a MACS-Quant VYB flow cytometer (Miltenyi Biotec) and FlowJo Software. Antibodies are listed in Supplementary Table S4 (**Table S4**). Specimens were analysed with confocal laser microscope (Leica TCS SP8). Images were acquired with Leica LAS AF software. MESP1-mCherry expressing cells were visualized in monolayer differentiation at day 3, using a Nuance™ Multispectral Imaging System.

RNA isolation, Quantitative Real-time PCR and Microarray Analysis

Total RNA was isolated using either the RNeasy micro (Qiagen) or Nucleospin RNA XS kits (Macherey Nagel) according to manufacturer's protocols, and treated with RNase-free DNase. Between 100 ng to 1 ug RNA was reverse transcribed into cDNA with I-scriptTM cDNA synthesis kit (Bio-rad), according to manufacturer's protocol. Real-time quantitative PCR was performed on a CFX96 Real Time System (Biorad). All experiments were carried out in triplicate with iTaq™ Universal SYBR® Green Supermix. Expression levels were normalised to the housekeeping gene hARP and multiplied by 1000 for the purpose of presentation. Primer sequences are listed in Supplementary Table S2 (**Table S2**). Microarray analysis was performed on three biologically independent replicates using Illumina human HT-12v4 arrays (ServiceXS). Illumina microarray data were processed using the lumi package and differentially expressed genes were identified using the limma package, included in R/Bioconductor[25, 26]. ConsensusPathDB-human (http://cpdb.molgen.mpg.de/) web server was used for gene ontology analysis. Additional analysis was performed using Genespring (Agilent Technologies).

RESULTS

Generation of a MESP1mCherry-NKX2-5eGFP dual cardiac reporter

To identify, purify and characterize human pre-cardiac mesodermal progenitors and study their further differentiation, we generated a dual fluorescent hESC reporter line by inserting the sequence encoding mCherry into the MESP1 locus of an existing NKX2-5^{eGFP} reporter line by homologous recombination **(Fig 1a, b)**. A PCR screening strategy was performed to identify correctly targeted clones **(Fig S1a,b)**. Twenty correctly targeted clones were obtained (11% targeting efficiency) from which four clones showed the loxP-flanked neomycin-resistance cassette was excised following transfection with Cre recombinase (not shown) [22]. We proceeded with clone 2.20B3 **(Fig. S1c)**. hESCs from clone 2.20B3 displayed a high immunoreactivity for NANOG, TRA-181, OCT-4, and EPCAM antibodies, confirming no loss of pluripotency during targeting and selection procedures **(Fig S2a,b)**. Furthermore, chromosome analysis demonstrated no abnormal karyotypes in these targeted lines **(Fig S2c)**.

To demonstrate that the targeted MESP1$^{Mcherry/w}$-NKX2-5$^{eGFP/w}$ hESC lines faithfully recapitulated MESP1 activity, hESCs were differentiated as monolayers to cardiac cells by addition of the growth factors BMP4 and Activin-A, as well as GSK-3-β inhibitor Chir99021 from day 0 to day 3, followed by the Wnt-signalling antagonist pathway inhibitor Xav939 from day 3 to day 7 of differentiation **(Fig 1c)**. Red fluorescence of mCherry was clearly visible in a subset of cells at day 3 of differentiation **(Fig 1d)**. Cells expressing mCherry appeared to have undergone EMT, based on their mesenchymal-like morphology.

MESP1mCherry-NKX2-5eGFP hESCs differentiate towards mesoderm and the cardiac lineage

Quantitative PCR (qPCR) analysis of the targeted cell line during differentiation towards the cardiac lineage confirmed transient expression of mCherry, peaking at day 3 of differentiation and mirroring endogenous *MESP1* expression **(Fig 2a)**. In agreement with previous *in vitro* and *in vivo* studies, expression levels of mesendodermal genes *Brachyury (BRY)* and *Eomesodermin (EOMES)* displayed peak values at day 2 of differentiation, before that of *MESP1*. Expression levels of *NKX2-5* were clearly detectable from day 9 of differentiation, just before the appearance of contractile areas. Expression of *NKX2-5* further increased at day 12, confirming the ability of the MESP1^{mCherry}-NKX2-5^{GFP} reporter line to differentiate towards the cardiac lineage. Flow cytometric analysis showed that MESP1-mCherry expressing cells first appeared around day 2, peaking at day 4 (8.0% ± 0.5% (mean ± SEM)), and becoming negligible by differentiation day 8 **(Fig 2b,c)**, reflecting the gene expression profile. NKX2-5-eGFP protein was detected later, first appearing on day 8 and continuing to increase to day 14 of differentiation (28.9% ± 1.7%), reflecting our previous findings[17]. These results demonstrated that the MESP1^{mCherry} and the NKX2-5^{eGFP} reporters allowed visualization of early MESP1-expressing progenitors, followed by the formation of NKX2-5-expressing cardiac cells during hESC differentiation.

Gene Expression Profiling of MESP1-mCherry progenitors

To further characterize the MESP1-mCherry progenitors, we isolated MESP1-mCherry positive cells (MESP1-mCherry pos) at day 3 of differentiation and compared their gene expression profile with MESP1-mCherry negative cells (MESP1-mCherry neg) and undifferentiated hESCs of the same line *(Fig 3a)*. A correlation heatmap between samples is shown in supplemental figure 3 *(Fig. S3)*. In three biological replicates, 722 genes showed an increase in expression of ≥ 1.5 fold when compared to MESP1-mCherry neg cells **(Fig 3b)**. In agreement with their mesodermal character, MESP1-mCherry pos cell fraction was enriched for EMT- and mesendodermal related genes, including *MESP1* itself, *SNAI1, SNAI2, MSX1, FOXC1, GSC, MIXL1, EOMES, LHX1, FOXF1, and IRX3.* Although the differences were greater in MESP1-mCherry pos cells, mesendodermal and EMT-related gene levels were higher in both MESP1-mCherry pos and neg fractions compared to undifferentiated hESCs, indicating that cells had undergone some form of mesoderm differentiation in the protocol used. Furthermore, the enrichment of N-cadherin (*CDH-2*) and vimentin (*VIM*), and decrease of E-cadherin (*CDH-1*) and *EPCAM* are also strongly associated with EMT. As expected, transcripts of pluripotency genes, including *SOX2, OCT3-4,* and *NANOG*, were significantly decreased. Furthermore, early cardiac genes, such as *GATA4, TBX3, MYL4, GATA6,* and *HAND1*, were enriched indicating the pre-cardiac character of MESP1-mCherry progenitors. In addition, genes encoding extracellular matrix proteins were increased, including several collagens, integrins, laminins, and fibronectin, as well as extracellular matrix (ECM)-receptors, such as integrin α-5 (*ITGA5*). These proteins could play an important role in stage-dependent processes such as mesoderm migration or niche formation of the pre-cardiac mesodermal cell population.

The microarray data also indicated the expression of genes of various signalling pathways in the MESP1-mCherry pos population, including Notch ligand *DLL3, BMP4, FGF17*, the BMP-antagonist *CER1*, Wnt inhibitors *DKK1* and *DKK4*, and other components of (non) canonical Wnt signalling, such as the ligand Wnt5a and (co)-receptors *FZD2*, *ROR-2*, and *RSPO3*. Protein complex analysis showed 100% enrichment of the non-canonical WNT5a/FZD2/ROR-2 complex (data not shown). In agreement with this, pathway analysis showed significant enrichment (P-value ≤ 0.01) in extracellular matrix organization, Wnt signalling pathway, cardiac progenitor differentiation, and heart development **(Fig 3c)**. qPCR analysis validated enhanced expression of a selection of the genes upregulated in the microarray data **(Fig 3d)**.

Surface Marker Expression Analysis of MESP1-mCherry progenitors

Further microarray analysis identified several genes encoding mesoderm associated cell surface proteins, including *NCAM1, ROR-2, CD13, CDH-2, CDH-11,* and *PDGFR-α*, which showed increased expression in MESP1-mCherry pos cells (Fig 3a). FACS analyses showed that large proportions of the MESP1-mCherry pos cells expressed cell surface proteins PDGFR-α (approximately 70%), CD13 (approximately 97%), and ROR-2 (approximately 96%) **(Fig 4a)**. However, none of these proteins marked the MESP1-mCherry pos population exclusively, since expression was also clearly detected in the MESP1-mCherry neg cell populations **(Fig 4b)**. Further, only a small percentage in both the MESP1-mCherry pos and neg population was co-expressing PDGFR-α and KDR, in agreement with the microarray analysis, which showed that KDR was not differentially expressed in the MESP1-mCherry pos population.

Generation of Cardiac Cells from Isolated MESP1-mCherry progenitors

To investigate whether MESP1-mCherry pos progenitors were able to differentiate towards the cardiac lineage, MESP1-mCherry pos cells were isolated on day 3 of differentiation, seeded as aggregates or in monolayer in the presence of Wnt-pathway agonist Xav939 until differentiation day 7, and maintained until differentiation day 14 **(Fig 5a)**. To evaluate how mesoderm progenitor populations are represented in MESP1-mCherry sorted derivatives, we initially performed FACS analysis on differentiation day 5 (2 days after replating). ROR-2+CD13+ cells were slightly enriched in the MESP1-mCherry pos population, compared to the MESP1-mCherry neg population; approximately 73% ± 9% (mean ± SEM) vs. 39% ± 16% respectively **(Fig 5b)**. Furthermore, almost all MESP1-mCherry derivatives on day 5 expressed PDGFR-α (approximately 98% ± 1% (mean ± SEM)), while only 50% ± 4% of the day 5 MESP1-mCherry neg derivatives did. Interestingly, there was no enrichment of a KDR+PDGFR-α+ cardiac progenitor population in the MESP1-mCherry pos population, compared to the MESP1-mCherry neg derivatives (approximately 22% ± 2% (mean ± SEM) vs. 28% ± 12%, respectively) **(Fig 5b)**. Similar results were found at other time points during differentiation (day 4-8) **(Fig. S4)**.

On day 8 and day 14 of differentiation, the expression of several cardiac-associated genes was enriched in the MESP1-mCherry pos derivatives compared to the MESP1-mCherry neg derivatives, including cardiac genes *NKX2-5* and *ISL1*, smooth muscle cell markers *SMA* and *MYH11*, and endothelial marker *PECAM-1* **(Fig 5c,d)**. Additionally, NKX2-5 expression, as determined by flow cytometric quantification of eGFP expression, was significantly enriched in the MESP1-mCherry pos population, treated with Xav939 (80% ± 4% (mean ± SEM)) when compared to MESP1-mCherry neg cells under similar differentiation conditions (18% ± 6%) **(Fig 6a,b,c)**. The eGFP+ cells detected in the MESP1-mCherry neg day 14 aggregates may have been derived from a small percentage of MESP1-mCherry pos cells that subsequently appeared in the MESP1-mCherry neg population following sorting. Indeed, low expression levels of mCherry and MESP1 could be detected in negative sorted samples at day 3 and day 4 (one day after reaggregation) **(Fig. S5)**.

The importance of inhibiting Wnt/β-catenin signalling following mesoderm induction to generate NKX2-5-expressing cardiac progenitors was highlighted by the low levels of eGFP in the untreated MESP1-mCherry pos (23% ± 4%) and MESP1-mCherry neg derivatives (0.50% ± 0.4%). Furthermore, in all conditions, the NKX2-5^{eGFP}+ derivatives were spontaneously contracting on differentiation day 14 (data not shown).

To determine whether other specific signalling pathways could affect cardiac differentiation from mesoderm, three other major pathways associated with cardiac differentiation were inhibited after MESP1-mCherry sorting and re-aggregation (from day 3 to day 7). Inhibitors of these pathways included TGF-β/act/Nodal antagonist SB431542, and BMP inhibitor Dorsomorphin. None of these inhibitors led to an increase in GFP expression at day 14 when compared to control treatment. In fact, GFP expression showed a tendency to decrease in the SB431542-treated cells **(Fig S6)**.

Day 14 MESP1-mCherry pos derivatives treated with Xav939, expressed various cardiac lineage markers, including α-Actinin (cardiomyocytes), PECAM-1 (endothelial cells), and SMA and SMC-myosin (expressed in immature cardiomyocytes as well as smooth muscle), indicating the MESP1-mCherry pos cell population can generate all three different cell types of the cardiac lineage **(Fig 6d)**. In addition, flow cytometric analysis showed enrichment of cTNT (approximately 51% ± 9% (mean ± SEM)) and PECAM-1 (approximately 2.2% ± 0.2% (mean ± SEM)) expressing cells in day 14 MESP1-mCherry pos derivatives, compared to the MESP1-mCherry neg derivatives (approximately 11% ± 5% and 0.62% ± 0.1% respectively) **(Fig 6e)**.

SIRPA+ VCAM1+ co-expression marks NKX2-5-eGFP+ cardiomyocytes

In addition to the characterization and differentiation of MESP1^{mCherry} progenitors and their further differentiation to NKX2-5eGFP cardiac cells, we studied whether the previous described cell surface proteins SIRPA and VCAM1 are specifically marking hESC-derived-cardiomyocytes[27, 28]. We analyzed SIRPA-VCAM1 and cTnT expression on cardiac cells at day 14 of differentiation of NKX2.5-eGFP-HESCs **(Fig S7)**. We showed that the SIRPA+VCAM1+ double positive population was highly enriched for NKX2-5-eGFP+ cells **(Fig S7a)**. In addition, the vast majority of NKX2-5-eGFP+ cells were cTnT+ **(Fig S7b)**. Similar results were described in a recent study by Skelton et al., in which SIRPA+VCAM1+NKX2-5-eGFP+ cardiomyocytes were shown to arise from a SIRPA+NKX2-5-eGFP+ intermediate progenitor population[28].

DISCUSSION

Genetically modified reporter lines are proving increasingly valuable for developing precise protocols for the directed differentiation of various lineages from human PSCs[29-32]. Here we increased the degree of complexity of the processes that can be analysed by creating a double transgenic reporter hESC line, taking advantage of endogenous activation of the key cardiac transcription factors *MESP1* and *NKX2-5*[17]. We introduced the fluorescent protein mCherry into the genomic locus of the mesodermal transcription factor MESP1, in *NKX2-5*^{eGFP} hESCs. As predicted from mouse developmental biology studies, the *MESP1*$^{\mathrm{mCherry}}$-*NKX2-5*$^{\mathrm{eGFP}}$ line firstly expressed mCherry under conditions of mesoderm induction, and later GFP+ cardiac progenitors and cardiomyocytes emerged in the cultures, demonstrating the value of hESCs as *in vitro* model of previously inaccessible stages of early human development.

MESP1-mCherry pos cells were first detected on day 2 of differentiation with peak levels between day 3 and 4, which is in agreement with our earlier observations during cardiac differentiation of hESCs[1]. We demonstrated that sorted and reaggregated MESP1-mCherry pos cells on differentiation day 3 have a much higher potential to differentiate to cardiovascular derivatives than MESP1-mCherry neg cells, as demonstrated by (i) the appearance of functional beating cardiomyocytes and (ii) high numbers of *NKX2-5*^{eGFP}+ cardiac cells in the population on day 14 of differentiation (11 days after sorting). At the same stage of differentiation, enriched gene and protein expression were found for PECAM-1, SMA, and cTNT, suggesting that MESP1-mCherry cells can differentiate to endothelial cells, smooth muscle cells and cardiomyocytes, respectively.

Microarray analysis of MESP1-mCherry pos and MESP1-mCherry neg cells revealed new insights into the transcriptional profile of human pre-cardiac mesodermal progenitors. In particular, transcription factors and signalling pathways associated with mesoderm formation, EMT, early cardiac development and components of the ECM were highly enriched in MESP1-mCherry pos cells. This is broadly in accordance with the molecular signature of MESP1-GFP expressing mouse cells[33]. However, disparities possibly related to interspecies differences, differentiation protocols, or technical variations (i.e. random integration of an isolated regulatory sequence of the *Mesp1* mouse promoter versus homologous recombination in the human *MESP1* locus) were evident. For example, the expression of cardiac transcription factors such as *HAND2, MEIS2, SMARCD3, HEY1, HEY2,* and *TBX2* was increased in MESP1-GFP positive cells, whereas in our study these genes were not enriched in the MESP1-mCherry pos population. However, expression of these genes was increased at later stages following cardiac differentiation of sorted cells, suggesting that human MESP1-mCherry pos cells may be a more restricted mesodermal population.

Transition of undifferentiated epithelial pluripotent stem cells to mesodermal cells requires the activation of EMT-associated genes. Indeed, MESP1-mCherry pos cells exhibit an obvious increase in levels of EMT-related transcription factors SNAI1 and SNAI2. In agreement, increased expression of other EMT-related genes, including Vimentin and CDH-2, and decreased levels of epithelial proteins, such as CDH-1 and EPCAM, suggested that EMT took place in MESP1-mCherry pos cells. In addition, enrichment of many ECM-related transcripts, such as those for collagens, integrins, fibrillin, matrix remodellers, and receptors, like the Apelin Receptor (APLNR), suggested formation of an appropriate niche for MESP1-expressing cells [34, 35]. Alterations in cell-cell and cell-matrix interactions are crucial for biological processes such as EMT, intracellular interactions, cell-fate decisions, or migration of pre-cardiac MESP1-progenitors from the anterior part of the primitive streak to form the bilaterally situated cardiac crescent *in vivo* [36-39]. Further study is required to identify the role of specific ECM components and receptors for expansion and further differentiation of MESP1+ cells.

Several signalling pathways were also significantly activated in MESP1-mCherry pos cells. In particular, components of the Wnt signalling pathway were affected, including the Wnt-inhibitors DKK1 and DKK4, the (non)canonical pathway ligands WNT4 and WNT5a, and the receptors ROR-2 and FZD2. Previous studies have indicated an important role for activated canonical Wnt-signalling during hESC to mesendoderm differentiation[40-43]. Subsequent inhibition of this pathway by DKK-1 resulted in improved cardiac differentiation, confirming the need to inhibit canonical Wnt-signalling following mesoderm formation for efficient cardiac differentiation[14, 40]. Conversely, at the same stage, activated non-canonical signalling has been associated with driving cardiac differentiation[44]. Non-canonical Wnt signalling has been linked with different molecular mechanisms, including canonical Wnt and/or β-catenin signalling inhibition[42, 45, 46], progenitor proliferation[47], and cellular movements and polarity. Here, we confirmed the importance of Wnt-inhibition in mesoderm to cardiac progression, as we found a major increase in the percentage of NKX2-5^{eGFP} expressing cells upon treatment with the chemical Wnt signalling antagonist, Xav939. Other pathways that have been shown to play crucial roles in the early stages of mesoderm formation, such as the BMP, Notch and FGF pathways, were also present in the MESP1-mCherry pos population, suggesting cross-talk of components of these pathways during these early pre-cardiac mesoderm stages[41].

We also examined expression of cell surface markers of mesodermal and cardiovascular progenitors during the differentiation of *MESP1*^{mCherry}-*NKX2-5*^{eGFP} cells. Several studies have reported that PDGFR-α and Flk1/KDR are expressed on early cardiovascular progenitor cells, giving rise to cardiomyocytes, endothelial cells, and vascular smooth muscle cells[5, 19, 20]. In recent studies using hESCs, an early mesoderm population co-expressing CD13 (ANPEP) and ROR-2 has been identified [18, 48]. This population consists of a KDR+PDGFR-α+ subpopulation, which has the potential to give rise to cardiovascular cells. In agreement, we demonstrated a significant increase in CD13, ROR-2, and PDGFR-α on day 3 MESP1-mCherry pos cells and on day 5 derivatives. However, these surface markers were also expressed in a high percentage of the MESP1-mCherry neg populations, indicating that they mark a broader population than only pre-cardiac mesoderm[27]. Interestingly, in MESP1-mCherry pos derivatives we did not find enrichment of KDR+ cells. These findings indicated that, besides the previously described Flk1/KDR+ PDGFR-α+ cardiac progenitor population, other MESP1-derived pre-cardiac mesoderm progenitor populations exist, that (co)express other (cell surface) markers, and have the potential to differentiate towards NKX2-5^{eGFP} expressing cardiac cells. In previous studies, the cardiac potential of other subpopulations besides the Flk1/KDR+ PDGFR-α+, was not tested[18, 20].

One of the major challenges in cardiovascular research is cardiac regeneration and improved cardiac function in heart failure patients. Tissue engineering, cellular therapy, endogenous activation of cardiac progenitors cells, or a combination of these approaches may lead to novel therapeutic strategies[49]. However, a key concern is the selection of the right cell type for transplantation or enhanced *in situ* proliferation, as this will be important for survival and engraftment, for safety (risk of tumor formation), and for the ability to form the required cell type(s) [50, 51]. A recent study showed the ability of a hESC-derived CD13+ROR-2+KDR+PDGFR-α+ cardiac progenitor population to migrate, couple, and form cardiac cells after being transplanted into ex vivo human fetal ventricular heart tissue, without formation of teratomas[18]. As pre-cardiac MESP1-progenitors are able to form almost all cell types of the heart, it will be of interest to examine both the developmental and regenerative capacity of these cells upon transplantation into both healthy and damaged hearts. Even though many hurdles have to be overcome, transplantation studies using cardiac-lineage-restricted progenitors may shed more light on future heart regeneration therapies. In addition, understanding the molecular cues that are important for expansion of mesodermal cells and their specific differentiation to the cardiac lineage may lead to novel therapies which are based on the progression of endogenous cardiac progenitor cells to functional specialized cardiovascular cells.

CONCLUSION

We conclude that the MESP1$^{mCherry/w}$-NKX2-5$^{eGFP/w}$ dual cardiac fluorescent reporter hESC line in this study is a useful tool for dissecting the molecular and cellular mechanisms related to the earliest steps in cardiac development in live cells. This is novel way to carry out fate analysis of this subpopulation. Purified mesodermal MESP1-mCherry pos cells efficiently differentiated to beating NKX2-5^{GFP}+ cells by inhibition of the Wnt-pathway. Besides a better understanding of underlying mechanisms, this human cardiac dual reporter line will make it feasible to identify and specify culture conditions for expansion of a mesodermal cell population and their further differentiation towards cardiac progenitor cells and specific subtypes of the cardiac lineage *in vitro*. This is of particular interest for future studies on cardiac toxicity, drug screening, tissue engineering, and regenerative medicine.

ACKNOWLEDGEMENTS

We thank Dorien Ward-Van Oostwaard for technical assistance; Kostas Gkatzis for help on statistical analysis using R/Bioconductor, the Department of Molecular Cell Biology, LUMC, for karyotyping analysis. S.C.D.H was supported by the Rembrandt Institute for Cardiovascular Disease. This study was supported by the Rembrandt Institute for Cardiovascular Science (S.H.), NHMRC, Australia (D.A.E.), ZonMw-MKMD-40-42600-98-036 (R.P.), ERC Advanced STEMCARDIOVASC-323182 (C.L.M.).

FIGURES

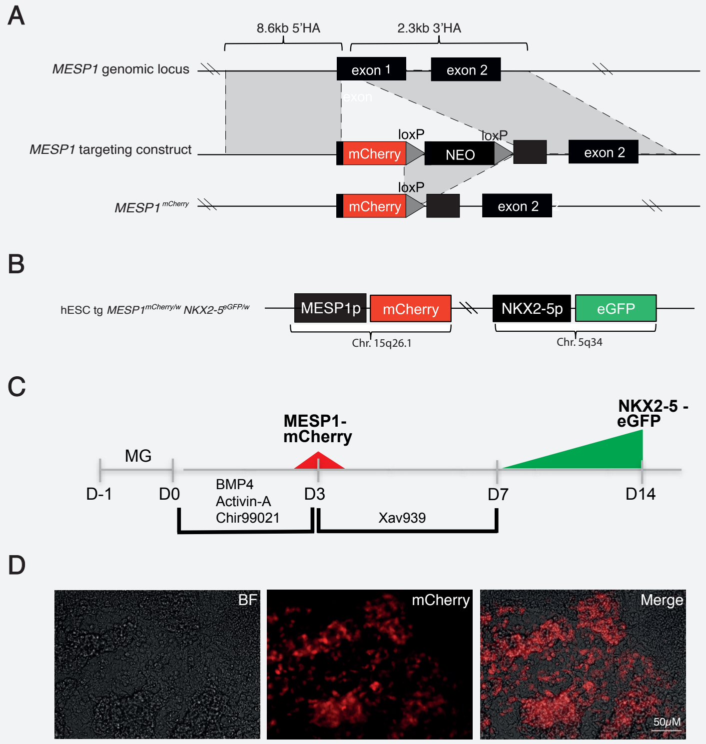


Figure 1. Generation of a MESP1mCherry-NKX2-5eGFP dual reporter hESC line. A: Schematic overview of wild-type MESP1 allele, MESP1 targeting construct, and targeted MESP1 allele. **B:** Schematic overview of the dual cardiac fluorescent reporter line. p: endogenous gene regulatory elements of either MESP1 (MESP1p) or NKX2-5 (NKX2-5p). **C:** Cardiac monolayer differentiation protocol in BPEL. MG: matrigel. Tg: transgenic. **D:** MESP1-mCherry expressing cells at day 3 of cardiac differentiation, visualized by a Nuance™ Multispectral Imaging System.

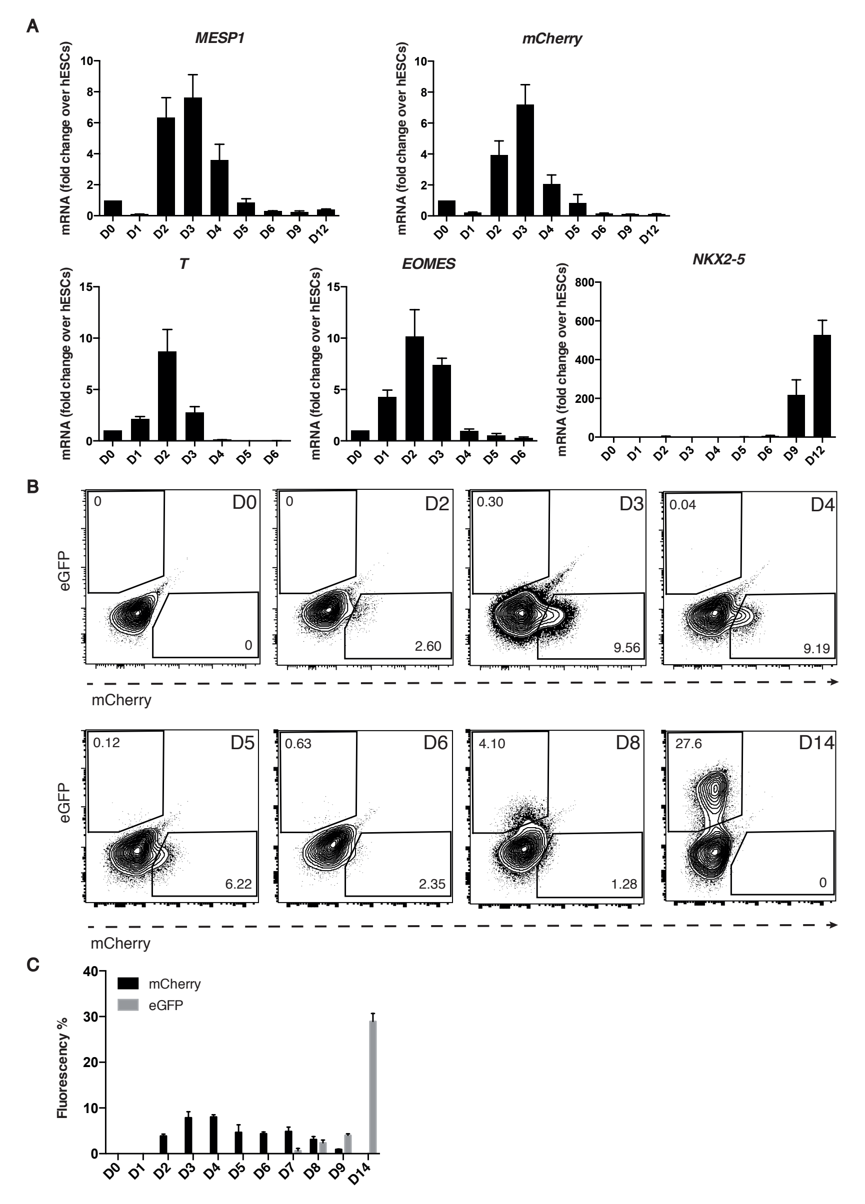
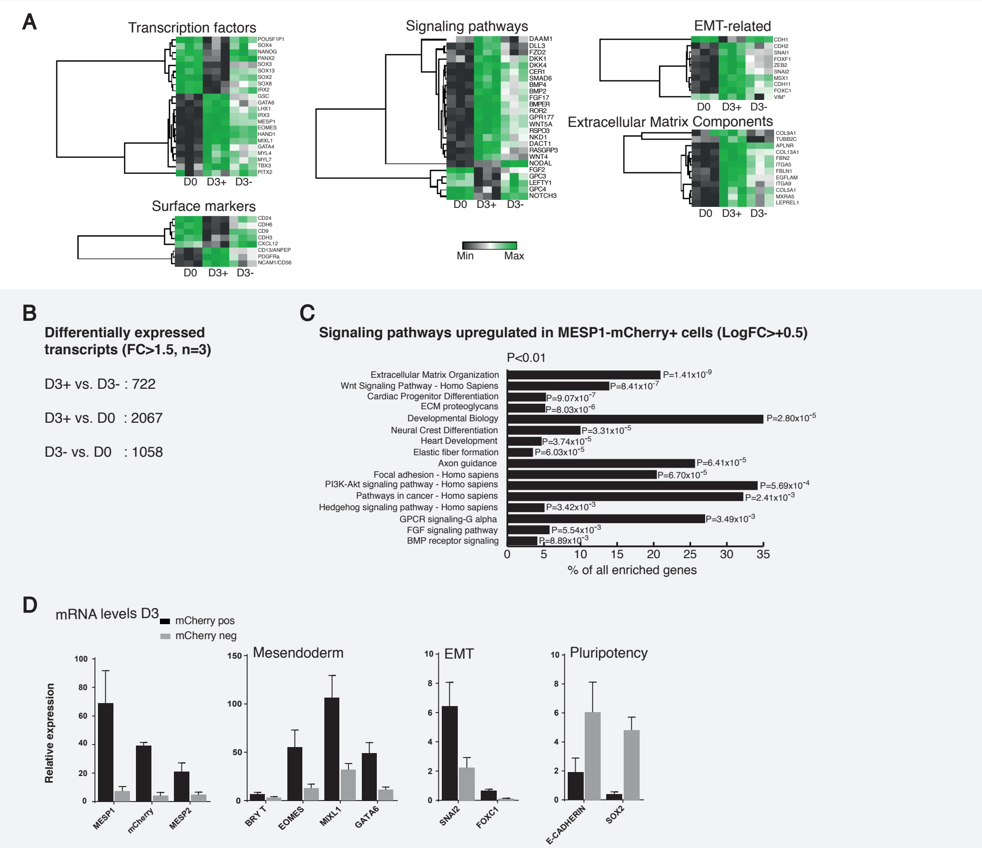


Figure 2. Timecourse of MESP1mCherry-NKX2-5eGFP hESCs differentiating towards the cardiac lineage. A: mRNA levels of mesendodermal genes *BRACHYURY T* (T) and *EOMESODERMIN (EOMES)*, pre-cardiac mesoderm gene *MESP1, mCherry* reporter gene, and early cardiac gene *NKX2-5* during monolayer differentiation. Results are normalized for the transcript expression at day 0 (undifferentiated hESCs) (n=3, error bars indicate SEM). **B:** Percentages of mCherry and eGFP protein expression upon monolayer differentiation. C: Quantification of percentages of mCherry and eGFP expressing cells during monolayer differentiation (n= 4-8; error bars indicate SEM).



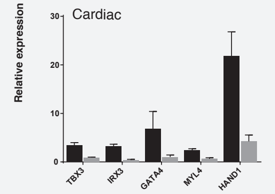


Figure 3. Characterization of MESP1-mCherry expressing cells at day 3 of monolayer differentiation.

A: Transcriptional profiling of three biological replicates of undifferentiated hESCs (D0), differentiation day 3 MESP1-mCherry pos cells (D3+), and MESP1-mCherry neg cells (D3-). D3+ samples were enriched for specific transcription factors, cell surface markers, ECM components, EMT-related genes, and signalling pathway ligands and receptors. Clustering of the samples is based on hierarchical clustering of microarray gene expression profiling. Shown are a selection of transcripts expressed in D3+ with a fold change ≥ 1.5 compared to D3- for three independent paired experiments (n=3).* indicates differences in only two independent experiments. GEO accession number: GSE56721. http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE56721 **B:** Number of transcripts that was ≥ 1.5 fold differentially expressed (n=3) between D3+ and D0, D3- and D0, and D3+ and D3- samples. **C:** Signalling pathway analysis of enriched genes in MESP1-mCherry expressing cells. P<0.01 for selected pathways. Percentages are based on all enriched genes. **D:** Validating microarray results by qRT-PCR analysis of selected transcripts in purified MESP1-mCherry pos and MESP1-mCherry neg cell populations at differentiation day 3 (n=3, error bars indicate SEM). Values are normalized for hARP levels.

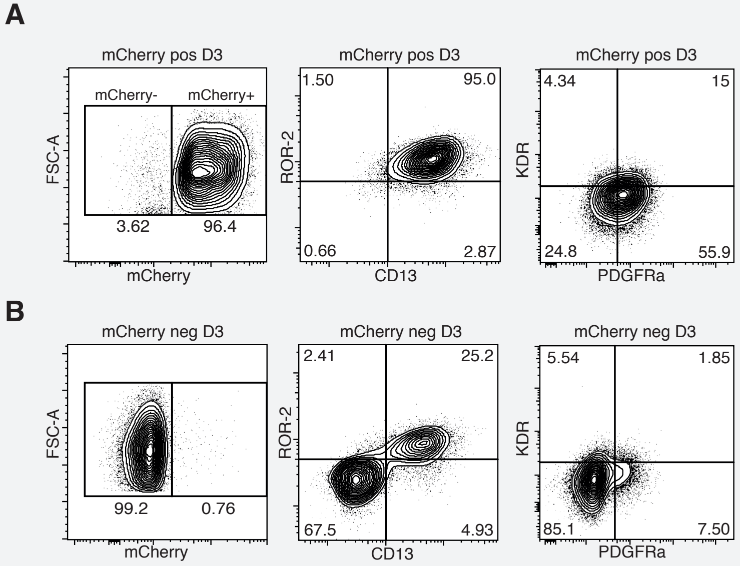


Figure 4. Flow cytometry analysis of cell surface markers CD13, ROR-2, KDR, and PDGFR-α on **(A)** MESP1-mCherry pos and **(B)** MESP1-mCherry neg sorted populations at day 3 of differentiation.

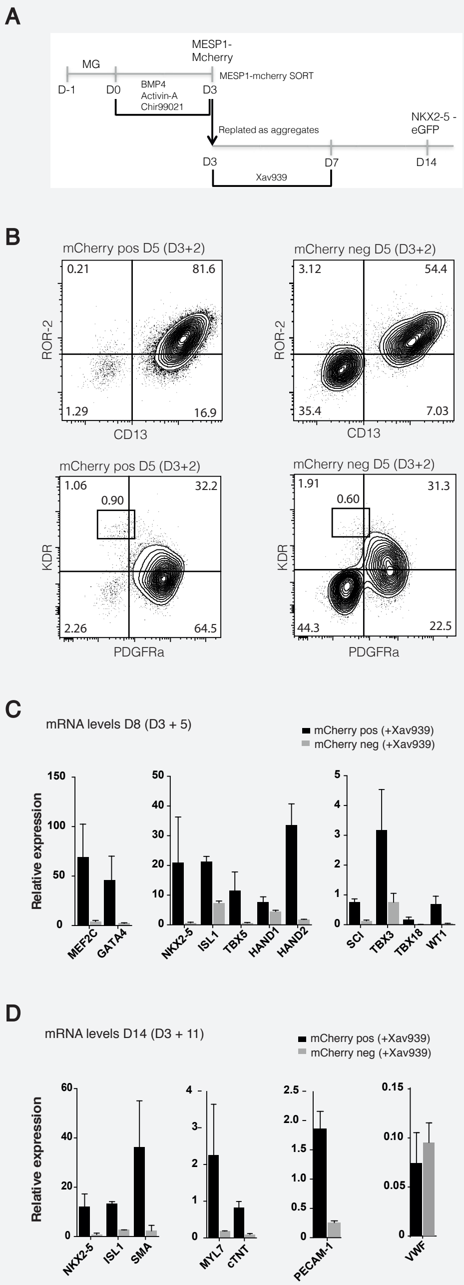


Figure 5. Cardiac differentiation induction from the MESP1-mCherry expressing pre-cardiac mesoderm progenitors. **A:** Schematic overview of the cardiac differentiation protocol for MESP1-mCherry cell fractions sorted by flow cytometry at day 3. **B:** Flow cytometric analysis of co-expression of cell surface proteins at differentiation day 5 (2 days after replating). **C,D:** Enrichment of mRNA transcripts of cardiac progenitor transcription factors at day 8 (n=3) and day 14 (n=2) in MESP1-mCherry pos and MESP1-mCherry neg derivatives (5 and 11 days after replating).

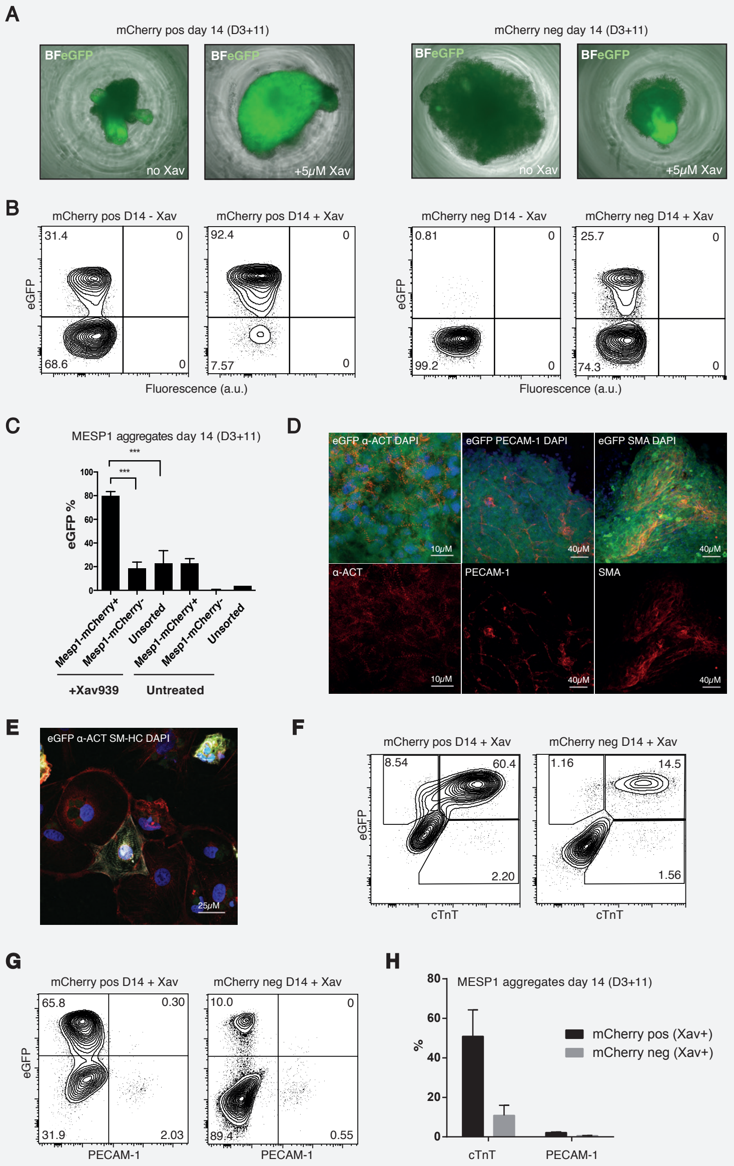


Figure 6. A: eGFP expression in MESP1-mCherry aggregates at day 14 of differentiation (11 days after replating), with and without treatment of Xav939. Brightfield (BF) overlay with eGFP. **B,C:** Percentage of cells expressing eGFP in day 14 MESP1-mCherry aggregates, with and without treatment of XAV939 (+sorted: n=7, -sorted: n=6, unsorted: n=2; error bars indicate SEM, ***p<0.001). **D:** Immunohistochemical stainings on day 14 MESP1-mCherry aggregates to analyse the presence of cardiac markers, including PECAM-1, SMA, SM-myosin, and α-actinin (α-ACT). **E:** Co-staining of α-ACT and SM-Heavy Chain (SM-HC) at d14 to discriminate smooth muscle cells from cardiomyocytes. **F,G,H:** Flow cytometric analysis of cTNT and PECAM-1 expression on day 14 MESP1-mCherry aggregates treated with XAV939 (n=2, error bars indicate SEM).

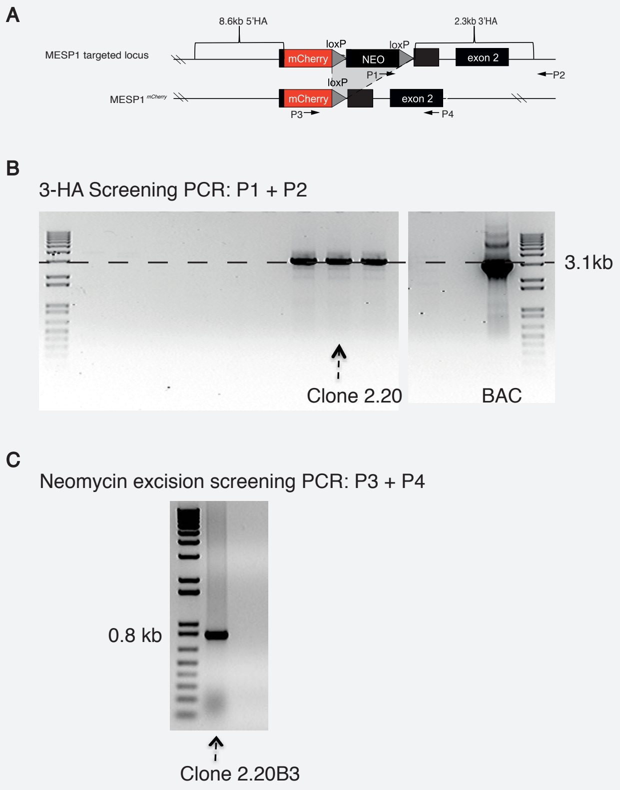


Figure S1. Genomic PCR experiments to confirm correct targeting of the MESP1 locus. A: Target screening PCR, using a forward primer annealing to the neomycin selection cassette and a reverse primer annealing outside the targeted region. Targeted clones show a PCR product of 3.1 kb. Clone 2.20 was used for neomycin cassette excision and further experiments. The modified Mesp1-mCherry bacterial artificial chromosome (BAC) was used as positive control. **B:** Neomycin excision of clone 2.20B3 was confirmed by a PCR reaction using a forward primer annealing within the mCherry cassette, and a reverse primer annealing within exon 2. Positively excised clones show a 800 bp PCR product.

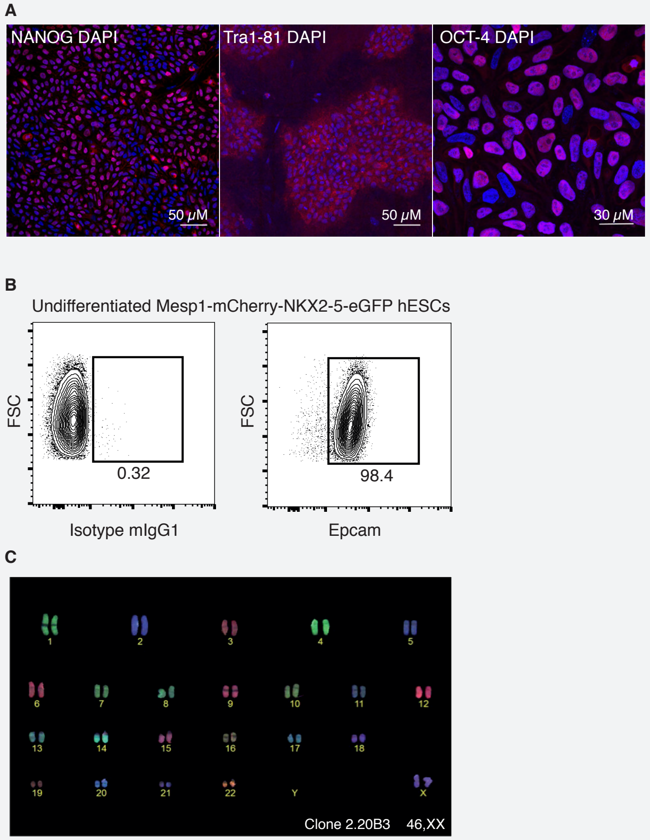


Figure S2. **A:** Immunohistochemical staining of MESP1^{mCherry}-NKX2-5^{eGFP} hESCs to confirm the expression of pluripotency markers after targeting. DAPI-stained hESCs (blue) with NANOG (red), OCT-4 (red), or TRA1-81 (red). **B:** Flow cytometric analysis of EPCAM expression on undifferentiated targeted hESCs. **C**: Karyotyping analysis of Clone 2.20B3 after targeting and neomycin cassette excision.

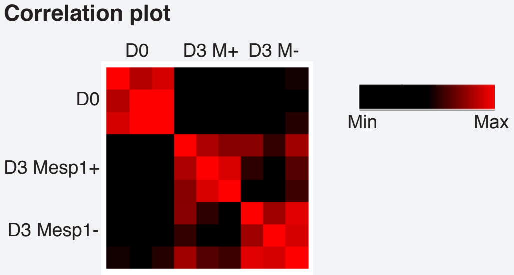


Figure S3. Additional information on MESP1-mCherry micro array experiments. **A:** Correlation heatmap between the three biological replicates of each timepoint: day 0 (D0), MESP1-mCherry positive cells (D3 M+), and MESP1-mCherry negative cells (D3 M-).

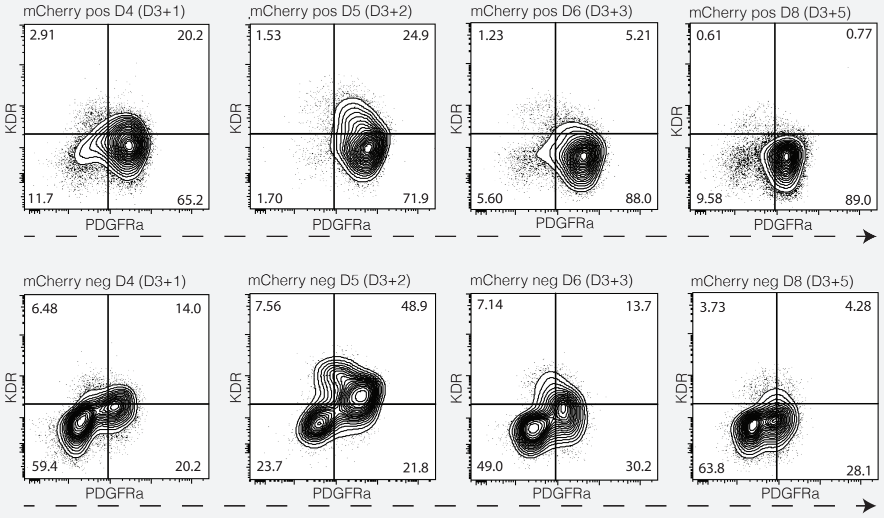
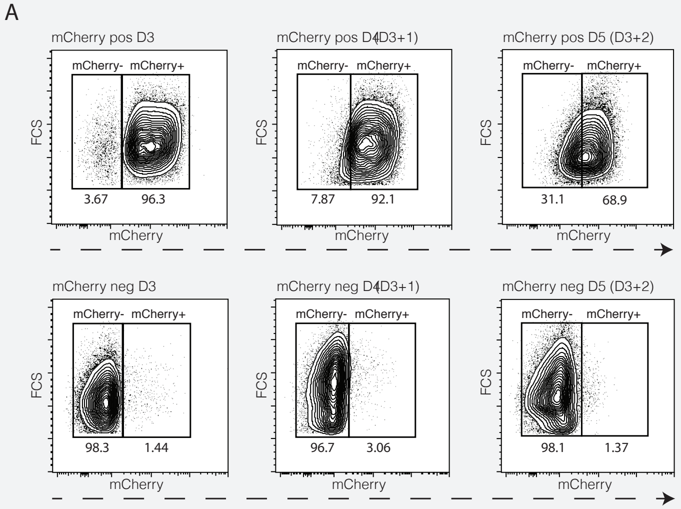


Figure S4. Expression profile of PDGFR-α and KDR during cardiac differentiation of MESP1-mCherry positive and negative sorted populations.



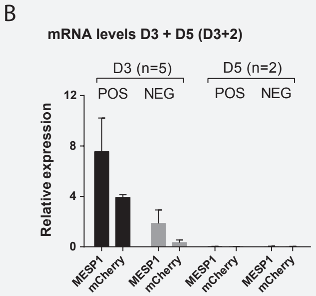


Figure S5. **A:** Re-analysis of mCherry expression on MESP1-mCherry pos and MESP1-mCherry neg sorted populations at day 3-5 of differentiation. **B:** mRNA expression levels of MESP1 and mCherry in MESP1-mCherry pos and MESP1-mCherry neg sorted populations at day 3 and 5 of differentiation.

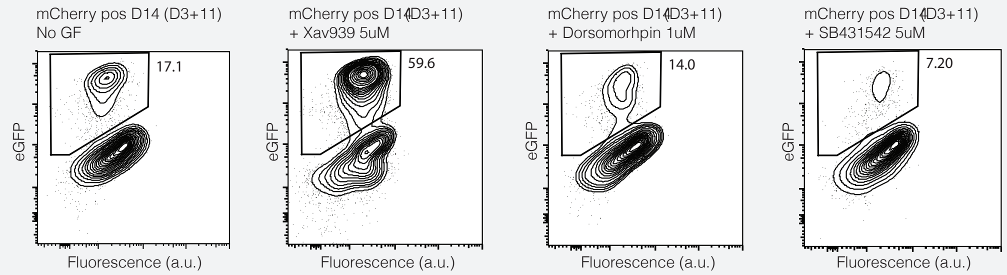


Figure S6. Effect of inhibition of specific signalling pathways after MESP1-mCherry sort and re-aggregation. NKX2-5-eGFP expression in MESP1-mCherry positive aggregates at day 14 of differentiation (11 days after replating). No GF: no addition of growth factors/inhibitors.

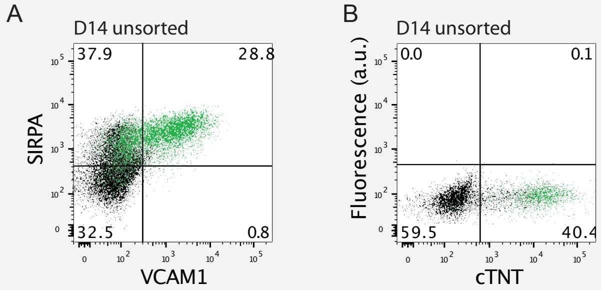


Figure S7. SIRPA+ VCAM1+ co-expression marks NKX2-5-eGFP+ cardiomyocytes. A: SIRPA+ VCAM1+ and **B:** intracellular cTnT expression on cardiac cells at day 14 of differentiation of NKX2.5$^{eGFP/w}$-hESCs. NKX2-5-eGFP+ cells are indicated in green.

Primer name	**Primer sequence**
MESP1_5arm_PvuI	5'GGAGAATTGCTTGAACCTGGGAGGTGGAGGCTGCAGTGAG CCGAGATCACGATCGCTCTCCTGAGTAGGACAAATC-3'
MESP1_3arm	3'CTTTAGTACATCCACAGAGTTGTGCAAGCATTGCCACTAATT TTAGAACGTTCCGCCTCAGAAGCCATAGAGC –5'
Target screen - F (P1)	5′- TCGCCTTCTTGACGAGTTCT -3′
Target screen - R (P2)	5′- AGGCAGCATGATGAGCCTTA -3′
Neo excision screen - F (P3)	5′- CCCCGTAATGCAGAAGAAGA -3′
Neo excision screen - R (P4)	5'-AGTTTCTCCCGCTCACTGG-3'

Supplementary table S1. Recombineering and screening primers.

Genes	**Forward primer**	**Reverse primer**
HARP	CACCATTGAAATCCTGAGTGATGT	TGACCAGCCCAAAGGAGAAG
MESP1	CTCTGTTGGAGACCTGGATG	CCTGCTTGCCTCAAAGTG
MCHERRY	CAAGTTGGACATCACCTCCCAC	ACTTGTACAGCTCGTCCATGC
MESP2	TCTCCAGAGCCCTGTCTGTC	GTCCCTGGTCCTCTGAGTTG
E-CADHERIN	ATTCTGATTCTGCTGCTCTTG	AGTCCTGGTCCTCTTCTCC
SOX2	AACCCCAAGATGCACAACTC	GCTTAGCCTCGTCGATGAAC
BRACHYURY T	ATCACCAGCCACTGCTTC	GGGTTCCTCCATCATCTCTT
EOMES	TCTGTCCCATTGAGCTTCTC	TCTGTCCCATTGAGCTTCTC
MIXL1	GGTACCCCGACATCCACTT	GAGACTTGGCACGCCTGT
HAND1	AATCCTCTTCTCGACTGGGC	TGAACTCAAGAAGGCGGATG
GATA6	TCCCCCACAACACAACCTAC	TGTAGAGCCCATCTTGACCC
SNAI2	CAGACCCTGGTTGCTTCAA	TGACCTGTCTGCAAATGCTC
FOXC1	CCTGCTTGCCTCAAAGTG	TCCTTCTCCTCCTTGTCCTT
MEF2C	CCACTGGCTCACCCTTCTCTGC	ACCCGTTCCCTGCACTAGTG
NKX2-5	TTCCCGCCGCCCCCGCCTTCTAT	CGCTCCGCGTTGTCCGCCTCTGT
ISL1	TGATGAAGCAACTCCAGCAG	GGACTGGCTACCATGCTGTT
GATA4	GACAATCTGGTTAGGGGAAGC	GAGAGATGCAGTGTGCTCGT
TBX5	ACATGGAGCTGCACAGAATG	TGCTGAAAGGACTGTGGTTG
SCI	AAGGGCACAGCATCTGTAGTCA	AAGTCTTCAGCAGAGGGTCACGTA
HAND2	ACATCGCCTACCTCATGGAC	TGGTTTTCTTGTCGTTGCTG
TBX3	CGGTTCCACATTGTAAGAGC	TTCGGGGAACAAGTATGTCC
TBX18	TTGCTAAAGGCTTCCGAGAC	AGGTGGAGGAACTTGCATTG
WT1	TGTGTGCTTACCCAGGCTGC	TGTGTGCTTACCCAGGCTGC
IRX3	GTGCCTTGGAAGTGGAGAAA	TAGGATGAGGAGAGAGCCGA
MYL4	TCAAAGAGGCCTTTTCATTG	CGTCTCAAAGTCCAGCATCT
CD45	ACTCTTGGCATTTGGCTTTG	CACTGGGCATCTTTGCTGTA
MYOGENIN4	CAGCTCCCTCAACCAGGAG	GCTGTGAGAGCTGCATTCG

Supplementary table S2. Quantitative PCR primers for human genes.

GO Accession	GO Term Name	P-value
GO:0048856	anatomical structure development	7,13E-19
GO:0048731	system development	3,51E-18
GO:0009653	anatomical structure morphogenesis	5,91E-18
GO:0007275	multicellular organismal development	7,23E-18
GO:0044767	single-organism developmental process	1,50E-17
GO:0009887	organ morphogenesis	4,15E-16
GO:0001501	skeletal system development	5,47E-16
GO:0007369	gastrulation	7,80E-16
GO:0044763	single-organism cellular process	1,75E-15
GO:0044707	single-multicellular organism process	7,82E-15
GO:0001704	formation of primary germ layer	1,47E-14
GO:0048598	embryonic morphogenesis	3,99E-14
GO:0048513	organ development	2,39E-13
GO:0048729	tissue morphogenesis	2,42E-13
GO:0009790	embryo development	2,63E-13
GO:0030154	cell differentiation	3,81E-13
GO:0050793	regulation of developmental process	4,65E-13
GO:2000026	regulation of multicellular organismal development	9,09E-13
GO:0009888	tissue development	1,02E-12
GO:0007155	cell adhesion	1,25E-12
GO:0003002	regionalization	2,33E-12
GO:0048869	cellular developmental process	3,11E-12
GO:0048568	embryonic organ development	4,28E-12
GO:0005578	proteinaceous extracellular matrix	5,21E-12
GO:0009952	anterior/posterior pattern specification	6,27E-12
GO:0072358	cardiovascular system development	1,17E-11
GO:0072359	circulatory system development	1,17E-11
GO:0048468	cell development	1,51E-11
GO:0048585	negative regulation of response to stimulus	1,74E-11
GO:0022603	regulation of anatomical structure morphogenesis	2,07E-11
GO:0045595	regulation of cell differentiation	2,22E-11
GO:0007498	mesoderm development	3,07E-11
GO:0032989	cellular component morphogenesis	3,11E-11
GO:0048646	anatomical structure formation involved in morphogenesis	3,58E-11
GO:0007154	cell communication	3,88E-11

Supplementary table S3. Gene Ontology classes overrepresented in day 3 MESP1-mCherry pos cells vs. day 3 MESP1-mCherry neg cells (LogFC>+0.5, P<0.01).

Cell surface antigens
Vioblue-conjugated-mouse-anti–EPCAM (Miltenyi Biotec, 130-098-092)
Mouse-anti-SSEA4 (Santa Cruz, SC59368)
BV421-conjugated-mouse-anti-PDGFRα (BD Biosciences, 562799)
BV421-conjugated-mouse-anti-CD13 (Biolegend, 301716)
A488-conjugated-mouse-anti-ROR-2 (R&D, FAB2064G)
PE-conjugated-mouse-anti-FLK1/VEGFR2/KDR (R&D Systems, FAB357P)
APC-conjugated-mouse-anti-PECAM-1 (eBiosciences, 17-0319)
APC-conjugated-anti-SIRPA (Miltenyi Biotec)
PE-conjugated-anti-VCAM1 (BD Biosciences)
Pluripotency markers
Mouse-anti-TRA-181 (Millipore, MAB4381)
Mouse-anti-OCT3/4 (Santa Cruz, sc5279)
Anti-NANOG (Abcam, Ab21603)
Cardiac markers
Rabbit-anti-smooth muscle actin (Abcam, Ab5694)
Mouse-anti-cardiac Troponin T (CT3, Developmental Studies Hybridoma Bank, University of Iowa)
Mouse-anti-PECAM-1 (Dako, JC70A)
Rabbit-anti-α-actinin (Abcam, Ab11007)
Rabbit-anti-myosin smooth muscle heavy chain (BTI Biomedical Technologies, BT-562)

Supplementary table S4. Antibodies.

SUPPLEMENTAL METHODS

Analysis of MESP1-mCherry Derivatives

MESP1-mCherry-sorted cells were re-analyzed on mCherry expression by flow cytometry on day 3 (immediately after sorting), 4 and 5 of differentiation (1 and 2 days after reaggregation of days sorted cells, respectively). Analysis on KDR and PDGFR-α expression on day 4, 6, and 8 of differentiation of MESP1-mCherry-sorted cells was performed as described before.
To measure RNA levels of mCherry and MESP1 at day 3 and day 5 of differentiation (0 and 2 days after sorting), RNA isolation and qRT- PCR were performed as described before.

Inhibition of specific signalling pathways after MESP1-mCherry sorting and re-aggregation.

MESP1-mCherry cells were differentiated, sorted, and re-aggregated as described before. Pathway inhibitors assessed were 5uM XAV939, 5uM SB431542 (#1614, R&D), and 1uM dorsomorphin (P5499, Sigma Aldrich), from day 3 until day 7. NKX2-5-eGFP expression was analysed by flow cytometry at day 14 of differentiation (11 days after sorting), as described before.

NKX2-5eGFP/w hESC Culture and Cardiac Differentiation

NKX2-5$^{eGFP/w}$ hESCs were cultured in E8 medium (Life Technologies) according to the manufacturer's protocol. Cells were differentiated in BPEL as described previously, with FACS analysis performed at d14, using the conjugated antibodies SIRPA-APC (Miltenyi Biotec) and VCAM1-PE (BD Biosciences). Flow cytometric analysis of Troponin T was performed as described previously.

REFERENCES

1. Beqqali A, Kloots J, Ward-van Oostwaard D, et al. Genome-Wide Transcriptional Profiling of Human Embryonic Stem Cells Differentiating to Cardiomyocytes. **STEM CELLS** 2006;24(8):1956–1967.
2. Keller G. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. **GENETICS** 2005.
3. Mummery CL, Zhang J, Ng ES, et al. Differentiation of Human Embryonic Stem Cells and Induced Pluripotent Stem Cells to Cardiomyocytes A Methods Overview. 2012.
4. Garry DJ, Olson EN. A common progenitor at the heart of development. **CELL** 2006;127(6):1101–1104.
5. Kattman SJ, Huber TL, Keller GM. Multipotent Flk-1+ Cardiovascular Progenitor Cells Give Rise to the Cardiomyocyte, Endothelial, and Vascular Smooth Muscle Lineages. **DEVELOPMENTAL CELL** 2006;11(5):723–732.
6. Moretti A, Caron L, Nakano A, et al. Multipotent Embryonic Isl1+ Progenitor Cells Lead to Cardiac, Smooth Muscle, and Endothelial Cell Diversification. **CELL** 2006;127(6):1151–1165.
7. Abu-Issa R, Kirby ML. Heart Field: From Mesoderm to Heart Tube. **ANNU. REV. CELL DEV. BIOL.** 2007;23(1):45–68.
8. Saga Y. MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube. DEVELOPMENT 1999:1–11.
9. Kitajima S, Takagi A, Inoue T, et al. MesP1 and MesP2 are essential for the development of cardiac mesoderm. **DEVELOPMENT** 2000;127(15):3215–3226.
10. Kitajima S, Miyagawa-Tomita S, Inoue T, et al. Mesp1-nonexpressing cells contribute to the ventricular cardiac conduction system. **DEV. DYN.** 2006;235(2):395–402.
11. Saga Y, Hata N, Kobayashi S, et al. MesP1: a novel basic helix-loop-helix protein expressed in the nascent mesodermal cells during mouse gastrulation. **DEVELOPMENT** 1996.
12. Saga Y. Mesp1 Expression Is the Earliest Sign of Cardiovascular Development. **TCM** 2001:1–8.
13. Lindsley RC, Gill JG, Murphy TL, et al. Mesp1 coordinately regulates cardiovascular fate restriction and epithelial-mesenchymal transition in differentiating **ESCs. CELL STEM CELL** 2008.
14. David R, Brenner C, Stieber J, et al. MesP1 drives vertebrate cardiovascular differentiation through Dkk-1-mediated blockade of Wnt-signalling. **NATURE CELL BIOLOGY** 2008;10(3):338–345.
15. Bondue A, Lapouge G, Paulissen C, et al. Mesp1 acts as a master regulator of multipotent cardiovascular progenitor specification. **CELL STEM CELL 2**008.
16. Chan SS-K, Shi X, Toyama A, et al. Mesp1 Patterns Mesoderm into Cardiac, Hematopoietic, or Skeletal Myogenic Progenitors in a Context-Dependent Manner. **CELL STEM CELL** 2013;12(5):587–601.
17. Elliott DA, Braam SR, Koutsis K, et al. NKX2-5eGFP/w hESCs for isolation of human cardiac progenitors and cardiomyocytes. **NAT METH** 2011;8(12):1037–1040.
18. Ardehali R, Ali SR, Inlay MA, et al. Prospective isolation of human embryonic stem cell-derived cardiovascular progenitors that integrate into human fetal heart tissue. **PROCEEDINGS OF THE …** 2013.
19. Yang L, Soonpaa MH, Adler ED, et al. Human cardiovascular progenitor cells develop from a KDR+ embryonic-stem-cell-derived population. **NATURE** 2008;453(7194):524–528.
20. Kattman SJ, Witty AD, Gagliardi M, et al. Stage-Specific Optimization of Activin/Nodal and BMP Signaling Promotes Cardiac Differentiation of Mouse and Human Pluripotent Stem Cell Lines. **CELL STEM CELL** 2011;8(2):228–240.

21. Costa M, Dottori M, Sourris K, et al. A method for genetic modification of human embryonic stem cells using electroporation. **NATURE PROTOCOLS** 2007.
22. Davis RP, Costa M, Grandela C, et al. A protocol for removal of antibiotic resistance cassettes from human embryonic stem cells genetically modified by homologous recombination or transgenesis**. NATURE PROTOCOLS** 2008.
23. Costa M, Sourris K, Hatzistavrou T, et al. Expansion of human embryonic stem cells in vitro. **CURRENT PROTOCOLS IN CELL …** 2008.
24. Ng ES, Davis R, Stanley EG, et al. A protocol describing the use of a recombinant protein-based, animal product-free medium (APEL) for human embryonic stem cell differentiation as spin embryoid bodies. NATURE PROTOCOLS 2008;3(5):768–776.
25. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. BIOINFORMATICS 2008.
26. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. **STAT APPL GENET MOL** BIOL 2004.
27. Dubois NC, Craft AM, Sharma P, et al. SIRPA is a specific cell-surface marker for isolating cardiomyocytes derived from human pluripotent stem cells. **NATURE BIOTECHNOLOGY** 2011;29(11):1011–1018.
28. Skelton R, Costa M, Anderson DJ, et al. SIRPA, VCAM1 and CD34 identify discrete lineages during early human cardiovascular development. **STEM CELL RESEARCH** 2014.
29. Micallef SJ, Li X, Schiesser JV, et al. INS GFP/w human embryonic stem cells facilitate isolation of in vitro derived insulin-producing cells. **DIABETOLOGIA** 2011;55(3):694–706.
30. Davis RP, Ng ES, Costa M, et al. Targeting a GFP reporter gene to the MIXL1 locus of human embryonic stem cells identifies human primitive streak-like cells and enables isolation of primitive hematopoietic precursors. **BLOOD** 2008;111(4):1876–1884.
31. Wang P, Rodriguez RT, Wang J, et al. Targeting SOX17 in Human Embryonic Stem Cells Creates Unique Strategies for Isolating and Analyzing Developing Endoderm. **CELL STEM CELL** 2011;8(3):335–346.
32. Cai Q, Bonfanti P, Sambathkumar R, et al. Prospectively isolated NGN3-expressing progenitors from human embryonic stem cells give rise to pancreatic endocrine cells. **STEM CELLS** 2014.
33. Bondue A, Tannler S, Chiapparo G, et al. Defining the earliest step of cardiovascular progenitor specification during embryonic stem cell differentiation. **THE JOURNAL OF CELL BIOLOGY** 2011;192(5):751–765.
34. Paskaradevan S, Scott IC. The Aplnr GPCR regulates myocardial progenitor development via a novel cell-non-autonomous, Gαi/o protein-independent pathway. **BIOLOGY OPEN** 2012.
35. Schenke-Layland K, Nsair A, Van Handel B, et al. Recapitulation of the embryonic cardiovascular progenitor cell niche. **BIOMATERIALS** 2011.
36. Cheng P, Andersen P, Hassel D, et al. Fibronectin mediates mesendodermal cell fate decisions. **DEVELOPMENT** 2013.
37. Hazeltine LB, Selekman JA, Palecek SP. Engineering the human pluripotent stem cell microenvironment to direct cell fate. **BIOTECHNOLOGY ADVANCES** 2013.
38. Yang JT, Rayburn H, Hynes RO. Embryonic mesodermal defects in alpha 5 integrin-deficient mice. **DEVELOPMENT** 1993.
39. Mahlapuu M, Ormestad M, Enerback S, et al. The forkhead transcription factor Foxf1 is required for differentiation of extra-embryonic and lateral plate mesoderm. **DEVELOPMENT** 2001.
40. Lindsley RC, Gill JG, Kyba M, et al. Canonical Wnt signaling is required for development of embryonic stem cell-derived mesoderm. **DEVELOPMENT** 2006.
41. Noseda M, Peterkin T, Simoes FC, et al. Cardiopoietic Factors: Extracellular Signals for Cardiac Lineage Commitment. C**IRCULATION RESEARCH** 2011;108(1):129–152.
42. Ueno S, Weidinger G, Osugi T, et al. Biphasic role for Wnt/β-catenin signaling in cardiac specification in zebrafish and embryonic stem cells**. PROCEEDINGS OF THE …** 2007.
43. Lian X, hsiao C, Wilson G, et al. Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling(2). **PROCEEDINGS OF THE …** 2012.
44. BRADE T, MANNER J, KUHL M. The role of Wnt signalling in cardiac development and tissue remodelling in the mature heart. **CARDIOVASCULAR RESEARCH** 2006;72(2):198–209.
45. Mikels AJ, Nusse R. Purified Wnt5a protein activates or inhibits β-catenin–TCF signaling depending on receptor context. **PLOS BIOLOGY** 2006.
46. Maye P. Multiple Mechanisms for Wnt11-mediated Repression of the Canonical Wnt Signaling Pathway. **JOURNAL OF BIOLOGICAL CHEMISTRY** 2004;279(23):24659–24665.
47. Povinelli BJ, Nemeth MJ. Wnt5a Regulates Hematopoietic Stem Cell Proliferation and Repopulation Through the Ryk Receptor. **STEM CELLS** 2014.
48. Tang C, Ardehali R, Rinkevich Y, et al. Isolation of primitive endoderm, mesoderm, vascular endothelial and trophoblast progenitors from human pluripotent stem cells. **NATURE BIOTECHNOLOGY** 2012;30(6):531–542.
49. Passier R, Van Laake LW, Mummery CL. Stem-cell-based therapy and lessons from the heart. **NATURE** 2008.
50. Buikema JW, Meer P, Sluijter J, et al. Concise Review: Engineering Myocardial Tissue: The Convergence of Stem Cells Biology and Tissue Engineering Technology. **STEM CELLS** 2013.
51. Hartogh den S, Sluijter J, Van Laake LW, et al. Chapter 11. Cellular therapy for the infarcted myocardium. In: Translational Cardiology. Springer; 2012.

CHAPTER 4

A comprehensive gene expression analysis at sequential stages of *in vitro* cardiac differentiation from isolated MESP1-expressing-mesoderm progenitors

Sabine C. den Hartogh[1], Katherine Wolstencroft[2], Christine L. Mummery[1],
Robert Passier[1,3]*.

Scientific Reports (Nature Publishing Group), *In Press, 2016*

1. Department of Anatomy and Embryology, Leiden University Medical Centre, Leiden, The Netherlands
2. Leiden Institute of Advanced Computer Science Leiden Institute of Advanced Computer Science, Leiden University, The Netherlands
3. Department of Applied Stem cell Technologies. MIRA Institute for Biomedical Technology and Technical Medicine. University of Twente, P.O.Box 217, Enschede, The Netherlands

ABSTRACT

In vitro cardiac differentiation of human pluripotent stem cells (hPSCs) closely recapitulates *in vivo* embryonic heart development, and therefore, provides an excellent model to study human cardiac development. We recently generated the dual cardiac fluorescent reporter MESP1$^{mCherry/w}$NKX2-5$^{eGFP/w}$ line in human embryonic stem cells (hESCs), allowing the visualization of pre-cardiac MESP1+ mesoderm and their further commitment towards the cardiac lineage, marked by activation of the cardiac transcription factor NKX2-5. Here, we performed a comprehensive whole genome based transcriptome analysis of MESP1-mCherry derived cardiac-committed cells. In addition to previously described cardiac-inducing signalling pathways, we identified novel transcriptional and signalling networks indicated by transient activation and interactive network analysis. Furthermore, we found a highly dynamic regulation of extracellular matrix components, suggesting the importance to create a versatile niche, adjusting to various stages of cardiac differentiation. Finally, we identified cell surface markers for cardiac progenitors, such as the Leucine-rich repeat-containing G-protein coupled receptor 4 (LGR4), belonging to the same subfamily of LGR5, and LGR6, established tissue/cancer stem cells markers. We provide a comprehensive gene expression analysis of cardiac derivatives from pre-cardiac MESP1-progenitors that will contribute to a better understanding of the key regulators, pathways and markers involved in human cardiac differentiation and development.

INTRODUCTION

HPSCs provide an excellent platform to model human heart development and cardiac differentiation *in vitro*. We and others have previously shown that there are close similarities between the temporal and sequential transcriptional activation of genes during human cardiac differentiation *in vitro* and in cardiac development *in vivo*[1,2]. Identification of key molecular events during cardiac differentiation from human pluripotent stem cells is instrumental for a better understanding of human cardiac development, and advancing the fields of regenerative medicine, disease modelling, and drug discovery. Although cardiac differentiation protocols have been improved significantly over the last years, in-depth knowledge of molecular mechanisms involved in cardiac lineage commitment is sparse. Specification of hPSCs towards the cardiac lineage is regulated by the precise temporal expression of transcriptional networks. Detailed information on these transcriptional networks and signalling molecules is essential for understanding the mechanisms underlying progression and expansion of progenitors and their differentiation to their specific subtype derivatives. When hPSCs differentiate towards the cardiac lineage, they first progress via an intermediate mesodermal stage, expressing transcription factor MESP1, before they further differentiate into cardiac derivatives, indicated by the increased expression of cardiac marker NKX2-5[3,4]. In order to understand how molecular mechanisms play a role in these lineage decisions, we previously generated the dual cardiac human embryonic stem cell (hESC) reporter line MESP1$^{mCherry/w}$NKX2-5$^{eGFP/w}$[4]. We characterised MESP1 progenitors by gene expression and surface marker analysis, and showed that MESP1 derivatives were predominantly enriched for NKX2-5 positive cardiomyocytes. Here, we have defined the temporal gene expression changes that occur during the differentiation of isolated MESP1-progenitors towards their cardiac derivatives. By classifying enriched genes according to similar expression patterns, functional pathway analysis, GO terms, sequence-specific DNA binding properties, and protein-protein interactions, we provide more insight on molecular signals and regulators involved in the sequential stages of mesoderm to cardiac differentiation. We confirm the enrichment of previously identified transcriptional regulators and cardiac functional genes, indicated by the predominant activation of components of the Wnt and TGFβ pathways. Interestingly, we identified novel cardiac networks, signalling pathways and markers, which may have an important role in the early cardiac differentiation and may be used as blueprint to evaluate or study human cardiac differentiation *in vitro*.

RESULTS

Collection of time points during cardiac differentiation

To obtain MESP1-mCherry progenitors for further differentiation, we differentiated and isolated MESP1$^{mCherry/w}$NKX2-5$^{eGFP/w}$ hESCs as described before[4]. Day 3 MESP1-mCherry positive and negative cells were replated as aggregates and further differentiated in the absence (control) or presence of Wnt-pathway inhibitor Xav939, which has been previously shown to enhance cardiomyocyte differentiation, **(Fig. 1a)**. Aggregates were collected 2 (D5), 4 (D7), 7 (D10) and 11 (D14) days after sorting and RNA was isolated subsequently. Experiments were performed in triplicates. To ensure efficient cardiac differentiations in the samples that were used for gene expression analysis, flow cytometry analysis was performed at day 14 of differentiation from a similar differentiation experiment **(Fig 1b)**. We found 52.3±10.7 % of NKX2-5-eGFP positive cardiac cells in the MESP1-mCherry positive derived population, treated with Xav939, which was, as expected, significantly higher than 4.8 ± 2.0 % in the MESP1-mCherry negative derived population, treated with Xav939 **(Fig. 1c).**

Global gene expression patterns segregate upon cardiac differentiation.

In order to study cardiac-specific enriched genes, we performed statistical analysis between replicates and generated gene lists based on a 1.5 fold change of expression level, with a gene-specific P-value<0.05, in the MESP1-mCherry positive population, treated with Xav939 (M+X+), compared to the MESP1-mCherry negative population, treated with Xav939 (M-X+) **(Fig 2a)**. With this, we selected genes that displayed both large changes as well as more subtle, but consistent, changes[5]. The number of differentially expressed genes increased upon further differentiation **(Fig. 2b)**, indicating a lower correlation between MESP1-positive and –negative samples upon further differentiation, and thus a more specialized gene expression pattern/phenotype. A top 100 list for both enriched and downregulated transcripts at each time point is given in **Supplementary Table S1**.

Gene Ontology analysis shows a clear shift from global development to heart development.

We determined which Gene Ontology (GO) categories were enriched for each specific differentiation stage, and could identify a clear shift between early cardiac progenitor stages D5 and D7, and the later cardiac-lineage-committed progenitors and cardiomyocytes (D10 and D14) **(Table 1)**. D7 GO terms were highly similar to those in D5, which included GOs that were more broadly developmental related. D10 and D14 were both highly enriched for heart-specific regulatory and functional GO terms, including heart development (BP00251), and muscle contraction (BP00173). To note, GO analysis of downregulated genes at each timepoint (FC>1.5 fold, P<0.05), did not show any tissue-specific terms, and is therefore not shown.

Transcriptional regulators during early cardiac commitment (D5)

Firstly, in order to show that sorted MESP1-mCherry progenitors leave the mesoderm stage upon further differentiation, we generated an expression heatmap up to day 14 for a selection of genes that were previously demonstrated to be enriched in day 3 MESP1-mCherry positive progenitors4 **(Fig. 3a).** As expected, the levels of mesoderm transcription factors *MESP1, MIXL1, Eomesodermin (EOMES),* and *Goosecoid (GSC)* show a transient expression with a peak at day 3. In order to identify genes that may play important roles in early cardiac differentiation development we selected genes that were upregulated (FC>1.5 fold, P<0.05) in the day 5 M+X+ population, when compared to the day 5 M-X+ control. From the 281 enriched transcripts, (potential) cardiac (co)-regulatory genes were selected based on their predicted transcriptional activity, DNA binding domains, and biological function **(Fig. 3b)**. Several transcription factors for which their role in early cardiac commitment has been shown previously could be identified based on their enrichment at day 5 of differentiation in the M+X+ samples (*GATA5, MEIS1, MEIS2, HEY1, IRX3, IRX5, GATA4*, and nuclear retinoic acid receptors *RARα* and *RARβ*, and *PPARγ*). Some of these genes maintained high levels of expression at later stages. In order to identify novel cardiac regulators, we focussed on proteins with conserved DNA binding motifs, such as homeodomain and zinc-finger domains motifs. This list of genes included *HOXB2, ZFPM1, ZMBTB16, ZNF503,* and *RUNX1T1*. To understand how these genes and their encoded proteins could be involved in networks related to early heart development, we performed analysis using the STRING database for interactomic connections with established key transcription factors (http://www.string-db.org/)[6] **(Fig. 3c)**. Using STRING, we predicted protein-protein associations based on *in vivo* and *in vitro* experimental assays, including gene co-occurrence in genomes (i.e. phylogeny), gene co-expression, gene fusion events, genomic neighbourhood (i.e. synteny), and experimental data such as co-immunoprecipitation and yeast two hybrid[6].

We found a high predicted interaction between MEIS1, MEIS2, PBX3, and HOXB2, based on binding complexes of MEIS proteins with other PBX and HOX homologs in drosophila and rodent models[7-9]. Moreover, studies have indicated a crucial role for MEIS1, MEIS2, PBX3 and HOXB2 in either heart development, including heart looping and chamber septation[2,10] or *in vitro* cardiac differentiation[2,11]. Interestingly, PBX3 has shown to induce either skeletal muscle in the presence of MyoD, a master regulator of skeletal muscle differentiation[12,13], or cardiac differentiation, in the presence of the cardiac transcription factor Hand2[12], indicating a crucial role for PBX3 as a cofactor during differentiation towards striated muscle. Moreover, MEIS1, MEIS2, HOXB2, and PBX3 were all upregulated upon Mesp1 induction in mouse ESCs, indicating that they act downstream of Mesp1[14].

The genes *ZFPM1* (FOG1; friend of GATA family-1), *ZBTB16*, and *ZNF503* belong to the class of zinc finger transcription factors. FOG1 contains nine zinc-finger domains and belongs to a family of proteins of which two genes have been identified in mammals: FOG1 and FOG2. FOG proteins interact with the N-terminal domain of GATA factors and modulate their activity[15] and have been shown to recruit nuclear receptor-transcriptional co-repressors and histone deacetylases (HDACs). Although the role of FOG1 in heart development is not well understood, one study in zebrafish showed the injection of an antisense morpholino directed against the homolog to murine FOG1 resulted in embryos with a large pericardial effusion and a deficient looping heart tube[16].

Another zinc-finger domain protein that we found highly enriched in MESP1-positive derivatives at day 5, and that is also upregulated upon Mesp1 induction in mESCs[14], is RUNX1T1 (runt-related transcription factor 1); a protein that is known to interact with transcription factors and to recruit a range of co-repressors to facilitate transcriptional repression[17]. In the human embryonic heart, RUNX1T1 expression is identified in both cardiomyocytes and endocardial cells[1,2,18]. Moreover, chromosome break points in the RUNX1T1 gene are associated with congenital heart disease[3,4,18]. Protein-protein interaction between RUNX1T1 and ZBTB16, a growth repressor in hematopoietic progenitor cells through its ability to recruit nuclear co-repressors such as histone deacetylases and Polycomb (PcG) family proteins, has been previously described[4,17] and was therefore also predicted following analysis in the STRING database **(Fig. 3c)**. Although no potential interactions in this cluster at day 5 were identified in the STRING database for Zinc Finger 503 (ZNF503), it has been previously classified as a potential human cardiac developmental regulator, based on its chromatin signature and its temporal expression level upon *in vitro* cardiac differentiation in hESCs[2,4].

Transcriptional regulators in early cardiac progenitors (D7)

Upon further differentiation of MESP1-derived cardiac committed cell lineages towards early cardiac progenitor stage (day 7 of differentiation), we found 660 differentially expressed genes (FC>1.5 fold, P<0.05) in day 7 M+X+ population, when compared to day 7 M-X+ control. Again, (potential) cardiac (co)-regulatory genes were selected based on their transcriptional activity, nucleotide binding domains, and biological function. We identified a subset of known and potential novel cardiac regulatory factors, including *GATA4, LHX2,* and *COUP-TF1 (NR2F1)*, and DNA-binding zinc-finger proteins *ZFPM2 (FOG2), TSHZ2,* and *ZFHX3* **(Fig. 4a)**.

STRING-based interaction analysis of genes enriched at day 7 showed predictive interactions of FOG2 with COUP-TF1 and GATA4 **(Fig. 4b)**. It has been previously shown that FOG2 represses COUP-TFII dependent synergistic activation of the atrial natriuretic factor promoter, suggesting that FOG2 functions as a co-repressor for both GATA and COUP-TF proteins[5,19,20], although it may also act as a co-activator in combination with other transcription factors[4,20,21]. Zinc finger homeodomain protein 3 (ZFHX3) contains 4 homeobox domains and 22 zinc finger domains, and is described as transcriptional repressor for myogenic differentiation through repression of the *MYF6* gene[6,22]. Based on chromatin signature and transient expression levels during *in vitro* cardiac differentiation in hESCs a regulatory role for ZFHX3 in human cardiac development has been suggested[2,6]. Moreover, sequence variants of ZFHX3 are associated with atrial fibrillation[7-9,23-25]. The zinc finger transcription factor TSHZ2, may be a potential transcriptional repressor for MEIS transcription factors **(Fig. 4b)**, based on interaction studies between conserved orthologs in drosophila. Its role in heart development is currently unknown.

Expression profile plots of potential regulatory genes show a distinct expression pattern throughout cardiac differentiation.

Stage-specific (co)-expression of transcriptional genes could indicate a regulatory role. Here, we show how similarities between expression profiles of PBX3, MEIS1, MEIS2, and HOXB2 throughout cardiac differentiation further strengthen the hypothesis that they act in a similar stage-specific molecular role, potentially as co-regulators **(Fig 5)**. In addition, similar expression profiles of GATA4 and FOG2 during cardiac differentiation support a potential co-regulatory role **(Fig. 5)**. Further, we show how expression levels of TSHZ2 distinctly increase throughout cardiac differentiation from day 5 onwards, suggesting a role for TSHZ2 in cardiac differentiation **(Fig. 5).**

Gene expression profiling of MESP1-derived cardiac cells (D10/D14)

Upregulated genes at day 10 (952) and day 14 (1062) of differentiated samples from M+X+ were classified as described before and expression levels were visualized in heatmaps **(Fig. 6a)**. Several genes from this list are known to play a key role in cardiac development, such as *NKX2-5, TBX18, WT1, TCF21, TBX2, HAND2, MEF2C, ISL1,* and *SMARCD3*. Interaction analysis of all transcription factors enriched in day 10 and day 14 samples showed a distinct regulatory cluster centred on NKX2-5 and GATA4 **(Fig. 6b)**. The presence of highly enriched levels of cardiac genes encoding for structural and sarcomeric proteins and ion channels indicates the presence of functional cardiomyocytes **(Fig. 6a)**. Moreover, while most structural genes are increased at cardiac progenitor stage day 7, levels of ventricular marker MYL2 could only be identified in clusters of upregulated genes at day 14. This is in agreement with previous studies, showing that MYL2 expression is only highly increased after approximately one month of cardiac differentiation[2,10,26], suggesting that MYL2 may serve as a marker for cardiomyocyte maturation.

Further, we could identify a node cluster of TCF21 and WT1, connected to TBX18; three regulatory transcription factors implicated in the formation of epicardial progenitors, indicating either the presence of these progenitors in our MESP1-derived cardiac aggregates or the potential to further differentiate to epicardial cells[2,11,27]. Furthermore, we identified enriched levels of growth factors VEGFA/B/C and receptor Neuropilin-1 (NRP1), ANGPT1, and surface marker CD34, indicating the derivation of vascular endothelial progenitors from MESP1 progenitors **(Fig. S1)**.

Merging transcriptional networks sequentially active throughout cardiac differentiation

In order to understand how cardiac transcriptional networks may be developed along cardiac lineage commitment and differentiation we generated a large interactive network, based on transcription factors that are enriched throughout all timepoints of differentiation (obtained through Gene Set Enrichment Analysis (GSEA) from Broad Institute) **(Fig. 7)**. Interestingly, we could identify the development of three distinct large node clusters, including retinoic acid nuclear receptors, containing PPARγ and RAR α/β, and two large interconnected cardiac networks centered on either GATA4 or NKX2-5.

Signalling pathways during cardiac differentiation

In order to study the role of specific signalling pathways during cardiac lineage differentiation, we performed pathway analysis using the KEGG pathway database source **(Fig. 8a, Supplementary Table S2)**) on the enriched genes at each stage of differentiation (M+X+ vs M-X+, FC>1.5 fold, $P<0.05$). From these results, we could clearly find an important role for Wnt-signalling pathway components during early cardiac lineage commitment (Day 5, Day 7), corroborating the findings from our previous study, in which we demonstrated non-canonical enrichment of Wnt pathway molecules, including Wnt5a and Wnt inhibitors DKK1 and DKK4 in MESP1 expressing progenitors[4,12,13]. Figure 8b shows an expression heatmap for enriched Wnt-pathway related genes across all time points **(Fig. 8b)**, including SFRP5, SFRP1, FZD2, FZD4, and WNT5A, TCF4, and EMT-regulators SNAI2, and LEF1 **(Fig. 8c)**. Interestingly, we identified significant enrichment of Wnt/β-catenin antagonists, including SFRP5, WNT5A, DACT1, and DACT3, in particular temporally peaking at early time points of cardiac differentiation, which would lead to downregulation of Wnt/β-catenin signalling, necessary for cardiac specification **(Fig. 8c)**[12,28]. Furthermore, through pathway analysis and protein-protein interaction analysis using STRING, we identified a prominent role for TGF-β signalling pathway components in early cardiac progenitor stage and cardiomyocytes (d7-d14) **(Fig. 8b, Supplementary Table S2)**. TGF-β signalling is important for cardiovascular development and plays an important role in epithelial-to-mesenchymal transition (EMT), in order to stimulate endocardial and epicardial transitions to mesenchymal cells of the heart[14,29]. Expression analysis of a selection of these transcripts demonstrates a similar profile, suggesting interaction and or co-regulation of Wnt/β-catenin antagonists, Wnt/β-catenin target element TCF4, and EMT-related transcription factors SNAI2 and LEF1, which can be activated by both Wnt/β-catenin and TGF-β signalling **(Fig. 8c)**[30].

Role of ECM proteins during cardiac differentiation

Extracellular matrix (ECM) proteins play an important role in the formation of the microenvironment for cells and provide different stimuli leading to activation of numerous molecular and cellular mechanisms, including migration, proliferation and differentiation of cells, as these events take place during the different stages of cardiac development[15,31,32]. Pathway analysis of enriched genes in the MESP1-derivatives indicated an important role of ECM signalling upon formation of the early cardiac progenitor populations and in established cardiomyocytes **(Fig. 8a)**. Interestingly, the composition of this cluster showed dynamic changes throughout cardiac differentiation **(Fig. 9a)**, suggesting the need for a specific composition of proteins forming a microenvironment or niche for cardiac progenitor cells and cardiomyocytes for optimal functioning in processes such as self-renewal, differentiation, and survival[16,32,33]. Moreover, we identified also other transiently upregulated levels of ECM components and proteases that have been implicated in heart developmental defects[14,32] **(Fig. 9a)**.

Surface marker expression analysis on MESP1-derived cardiac progenitors

An efficient approach to isolate specific subpopulations for translational applications will be the use of antibodies against distinct cell surface markers. For this, we analysed expression patterns of a large variety of cell surface markers that were enriched in the MESP1-derived cardiac populations at the different time points **(Fig. 9b)**. In our previous study we found that cell surface markers N-cadherin (CDH2), CD13 (ANPEP), and ROR[24,17] are broadly expressed in mesodermal cells and not specifically for MESP1-expressing progenitors. Surface proteins that show a transient peak expression at day 5 (M+X+), include TMEM88, CD82, TMEM171, LGR4, and CD74. From these, only TMEM88 has been clearly associated with heart development and acts downstream of GATA factors in the pre-cardiac mesoderm to specify lineage commitment of cardiomyocyte development through inhibition of Wnt/β-catenin signaling[34,35]. Interestingly, LGR4 belongs to the Leucine-rich Repeat-containing G protein-coupled receptor family with a close relation to LGR5 and 6, known stem cell marker for different tissues. LGR4 is expressed in the heart[36,37], although not exclusively, but no role has been described so far for heart development or cardiac stem cells. Surface markers that show high enrichment at day 5 of differentiation, with a continuous enrichment upon further cardiac differentiation (up to final cardiomyocyte formation) include GPC2, FZD4, PDFGRα, NCAM1, FLRT2, TMEM66, and GPR124. We, and others, have shown that PDFGRα and NCAM1 broadly mark mesoderm progenitors[4,38]. In early embryonic development, Fzd4 is highly expressed in the cardiac crescent, head mesenchyme, and later in the developing heart tube[39]. Loss of Fzd4 shows a decrease in density and branching of small arteries in the developing mouse heart[40]. Flrt2 and Gpr124 could both be potential markers for cardiac lineage-commitment. Flrt2 is abundantly expressed in developing mouse heart tissue and loss of Flrt2 results in in impaired expansion of the compact ventricular myocardium[41]. G-protein coupled receptor GPR124 shows a similar stage-specific expression pattern as *FLRT2*. However, *GPR124* is expressed on endothelial cells during blood vessel development, but not specifically expressed in the heart[42]. Transmembrane protein 66 (TMEM66) functions as calcium ion transmembrane transporter, and is enriched in MESP1 progenitors and their cardiac derivatives up to day 14. However, *in vivo* it has been shown that TMEM66 is not specific for the developing heart[43].

Surface protein transcripts that are specifically enriched in late cardiac progenitors (day 10), and cardiomyocytes (day 14) include *TMEM151A, VCAM1, TMEM71, TMEM173*, and *SIRPA*. VCAM1 and SIRPA have been previously identified to be specific human CM markers[3,44,45]. Mouse embryo expression databases describe TMEM71 in mesoderm tissues, including atria and ventricles. Not much is known about TMEM151A and TMEM173 expression in the heart.

DISCUSSION

For studying the temporal expression of key molecules involved in cardiac lineage commitment upon differentiation, we recently generated the dual cardiac fluorescent reporter MESP1$^{mCherry/w}$NKX2-5$^{eGFP/w}$ hESC line, allowing us to isolate early pre-cardiac mesoderm progenitors and follow their further differentiation towards NKX2-5 expressing cardiomyocytes[4]. In the current study, we used this dual cardiac reporter line to perform molecular profiling during cardiomyocyte differentiation from MESP1-expressing pre-cardiac progenitors by genome-wide transcriptomic analysis encompassing crucial events during cardiac differentiation from hPSCs: early cardiac lineage commitment, early cardiac progenitor stage, late cardiac progenitor stage, and functional cardiomyocytes. Global gene expression analysis and gene ontology analysis confirmed and visualized that expression patterns became more specified towards cardiac differentiation, shifting from developmental networks to cardiovascular-specific developmental networks. Throughout differentiation, we identified sequential expression patterns of key cardiac transcription factors, followed by enriched levels of structural and functional cardiac genes at day 10 and day 14 of differentiation, indicating that in the presence of Wnt-pathway inhibitor Xav939, isolated MESP1 progenitors have the preference to differentiate to the cardiac lineage. Using pathway analysis we identified a prominent role for Wnt-pathway antagonists during early cardiac lineage commitment leading to inhibition of Wnt/β-catenin signalling. Furthermore, using STRING software[6] that builds functional protein-association networks based on compiled available experimental evidence, we identified a subset of transcription factors during early lineage commitment: *HOXB2, ZFPM1, ZBTB16, ZNF503, RUNX1T1, ZFPM2,* and *ZFHX3*. Zinc-finger proteins ZFPM1 (FOG1) and ZFPM2 (FOG2) belong the FOG (friend of GATA) gene family, however, their role in heart development has not been elucidated yet. Based on previous studies, FOG proteins interact with GATA and COUPTF proteins, key transcription factors for early developmental processes, including heart development and lineage commitment[19-21]. FOG1 and FOG2 are indicated as co-activators or co-repressors of GATA- and COUPTF-activity on downstream cardiac genes, dependent on the transcriptional network that is active, pointing towards a putative role in cardiac lineage commitment. Besides their possible role in cardiac differentiation, FOG1 is expressed in blood islands of the yolk sac in mice and act as a cofactor with GATA1 to induce transcriptional activity of downstream genes in erythroid and megakaryocytic cell differentiation[46,47], which may indicate the presence of MESP1-derived hematopoietic progenitors in our cultures[48,49]. Interestingly, FOG1 antagonizes GATA-1 activities in other cell lineages[46]. Repression of other lineages may also lead to a preferred induction of the cardiac lineage. Here, we did not found an enrichment of GATA-1, nor did we find enrichment of surface receptor KDR, or ETV2 and Tal1, transcription factors that are important for a MESP1-progenitor derived hematopoietic lineage[48]. In contrast, we did find increased expression of GATA4, 5 and 6 at this stage of development, which could indicate a potential binding of FOG1 to the conserved amino zinc finger of these factors[15,20].

Furthermore, homeodomain protein HOXB2 has been previously identified as potential cardiac regulator, based on its epigenetic signature[2], although a more extensively investigated role for HOXB2 in anterior-to-posterior patterning has been described in hindbrain development[7,11]. In our study, we found temporal stage-specific co-expression of HOXB2 with PBX3, MEIS1 and MEIS2. STRING software showed binding evidence between these genes, based on experimental evidence of heterodimeric transcription complexes between PBX, HOX, and MEIS family genes[7]. Although a role for this predicted complex has not yet been described in the heart, both PBX3 and MEIS1 and MEIS2 have been implicated in heart development.

Network analysis revealed another cluster of transcription factors, consisting of RUNX1T1 and ZBTB16, which were first identified at day 5 of differentiation but maintained high expression levels throughout the course of cardiac differentiation. RUNX1T1 is described as a co-repressor for ZBTB16, important for hematopoietic lineage differentiation[17]. It is not clear whether these factors may also have a role during early cardiac differentiation, since increased expression is maintained throughout cardiac differentiation, or whether absence of other key regulators of hematopoietic differentiation prevented a hematopoietic molecular profile. Similarly, zinc finger transcription factor ZNF503, described during hindlimb and brain development, may also have a role in cardiac differentiation, since peak expression was observed at day 5 with continued expression levels throughout differentiation. ZFHX3 is another zinc finger homeobox domain containing protein that finds enriched levels in cardiac progenitors and cardiomyocytes. SNP variants in the coding sequence have been correlated to atrial fibrillation, suggesting a role in heart development[23,24].

Furthermore, we studied the temporal expression patterns of enrichment cell surface markers that could be useful for efficient cardiac progenitor and/or cardiomyocyte isolation experiments. Several cell surface markers have been identified including LGR4, which is an R-spondin receptor with strong positive effect on Wnt signalling, and could play a role in self-renewing capacity of early cardiac progenitors[50,51]. Its homologs LGR5 and LGR6 are well-known stem-cell-growth markers in other organ stem cells, including that of the intestine, stomach, and hair-follicle[52], which makes the role and expression of LGR4 in cardiac progenitors of high interest to study. Furthermore, we identified other cell surface markers, and found that transmembrane protein transcripts *TMEM151A*, *TMEM71,* and *TMEM173* show a stage-specific enrichment similar to that of *SIRPA* and *VCAM1*, both human-specific cardiomyocyte markers[44].

Another key finding in this paper is the importance of ECM components upon cardiac differentiation. For example, we find transcripts as COL9A2 and COL23A1 temporal-specific enriched at day 5 of differentiation, where a large number of other collagen transcripts are enriched throughout complete cardiac differentiation. Similar stage-specific patterns are seen for other ECM proteins, such as Laminin gamma-3 subunit (LAMC3, enriched at day 3), and Laminin alpha-2 subunit (LAMA2, enriched at day 14).

A complete understanding of how ECM proteins are involved in heart development is lacking. Increasing our knowledge of stage-specific ECM-controlled steps in early heart development will be valuable for understanding pathology of diseased hearts.

Furthermore, the variety of enriched cardiovascular transcripts in day 10 and day 14 M+X+ samples, including HAND2, which is expressed throughout the heart predominantly in the right ventricle; epicardial progenitor transcription factors WT1, TCF21, and TBX18; and vascular endothelial progenitor surface marker CD34, vascular growth factors VEGFA/B/C, and cell surface receptor tyrosine kinase (TEK/TIE2), which is required for normal angiogenesis and heart development during embryogenesis, indicate the mixture of cell populations that derives from MESP1 progenitors, or the requirement of these transcripts for optimal cardiomyocyte differentiation cultures.

To conclude, results from our comprehensive gene expression analysis revealed several potential novel cardiac regulators, either as (co)-activator, or (co)-repressor, dependent on the transcriptional network active. Future studies will clarify the role of these identified factors during early cardiac differentiation and whether they have multiple roles during lineage. In addition, besides their role in lineage specification, these factors may also play additional roles in other cellular processes such as proliferation, differentiation, or cell survival. Therefore, to further validate the role of predicted regulatory genes in early human heart development and/or in the onset of congenital heart disease, experimental interaction analysis, direct downstream gene target analysis, and knockdown studies *in vivo* and *in vitro* will be of high interest.

METHODS

hESC maintenance, cardiac differentiation, and sample collection

MESP1$^{mcherry/w}$NKX2-5$^{eGFP/w}$ hESCs were cultured and differentiated as described before[4]. In brief, hESCs were induced with BMP4 (30 ng/mL, R&D Systems), Activin A (20 ng/mL, Miltenyi Biotec), and Chir99201 (1.5 µM, Axon Medchem). At day 3 of differentiation, MESP1-mCherry positive and negative progenitors were isolated by FACS and aggregated for further differentiation, in the presence of 5 µM Xav939 (R&D). Total RNA was isolated using Nucleospin RNA XS kits (Macherey Nagel), at day 5 (2 days after replating), day 7, day 10, and day 14 of differentiation **(Fig. 1)**. At day 14 of differentiation, samples were analysed on NKX2-5-eGFP expression by flow cytometry, as described before[4].

Gene expression micro array and data analysis

RNA quality control, RNA labeling, hybridization and data extraction were performed at ServiceXS B.V. Microarray analysis was performed on three biologically independent replicates using Illumina human HT-12v4 arrays (ServiceXS B.V.). Data analysis was performed using Genespring (Agilent Technologies). Previously published micro array data that was complementary used in this study is numbered as: GSE56721. First, to filter probe sets on outlying values, we performed a one-ANOVA significant analysis test. P-values were corrected using the Benjamin-Hochberg method (corrected P-value<0.05). In order to select enriched genes in MESP1-mCherry positive derivatives compared to MESP1-mCherry negative derivatives, we performed statistical analysis, using a moderated T-test, on normalized values from three biological replicates. P-values were corrected using the Benjamin-Hochberg method. By using a volcano plot view, enriched genes were selected by FC>=1.5 fold value difference in MESP1-mCherry positive derivatives compared to MESP1-mCherry negative derivatives, with a P-value < 0.05 **(Fig. 2)**. GO analysis was performed on the selected genes, with a multiple-GO-term correction using the Benjamini-Yekutieli method, and a P-value cut-off of P<0.05. For GO analysis we used the DAVID Bioinformatics Recources 6.7 database from NIH (https://david.ncifcrf.gov/). Pathway analysis on the FC>=1.5 fold selected genes was performed using KEGG pathway databases. Pathways with a P-value cut-off of P<0.05 were selected. Heatmaps of selections of enriched genes were generated with Gene-E (http://www.broadinstitute.org/cancer/software/GENE-E/index.html). Genes were hierarchical clustered through the one minus Pearson correlation. In order to be able to predict temporal gene networks based on our selection of enriched genes in MESP1-positive derivatives, we screened genes on DNA binding domains and transcription factor activity (using Gene Set Enrichment Analysis (GSEA) at www.broadinstitute.org), and used the STRING database to integrate protein-protein interactions. This online available database provides a prediction pipeline for inferring protein-protein associations, covering more than 2000 organisms, with scalable algorithms for transferring interaction information between organisms, based on prediction, and *in vivo* and *in vitro* experimental assays, including gene co-occurrence in genomes (i.e. phylogeny), gene co-expression, gene fusion events, genomic neighbourhood (i.e. synteny), text mining, and experimental data such as co-immunoprecipitation and yeast two hybrid[6]. STRING extract experimental data from BIND, DIP, GRID, HPRD, IntAct, MINT, and PID databases. Cluster positions in the network are determined by an algorithm that is based on global confidence binding score (medium >0.4 or high>0.7). Based on this score, clustering of gene nodes (visualized by node colours) was determined by applying the Markov Cluster Algorithm[53]. Thus, interacting proteins with a higher global score have more chance to end up in the same cluster.

In order to understand how developmental networks are build up throughout cardiac differentiation, we extracted STRING data into Cytoscape[54] and generated a large interrogating network from protein-protein interactions from day 5, day 7, and day 10 (interactions with a high confidence > 0.7 are visualized).

ACKNOWLEDGEMENTS

We thank Kostas Gkatzis for his preliminary help on statistical analysis using R/Bioconductor. S.C.D.H. was supported by the Rembrandt Institute for Cardiovascular Science. R.P. by ZonMw-MKMD-40-42600-98-036. C.L.M. by ERC Advanced STEMCARDIOVASC-323182.

Author Contributions Statement

R.P. and S.C.D.H. designed the study. S.C.D.H conducted the study, and wrote the manuscript. K.W. contributed to conduction of the study. C.L.M., R.P. and K.W. reviewed the manuscript.

Competing financial interests

R.P. and C.L.M. are members of the scientific advisory board of Pluriomics B.V, but declare to provide an unbiased article. S.C.D.H. and K.W. declare no potential conflict of interest.

FIGURES

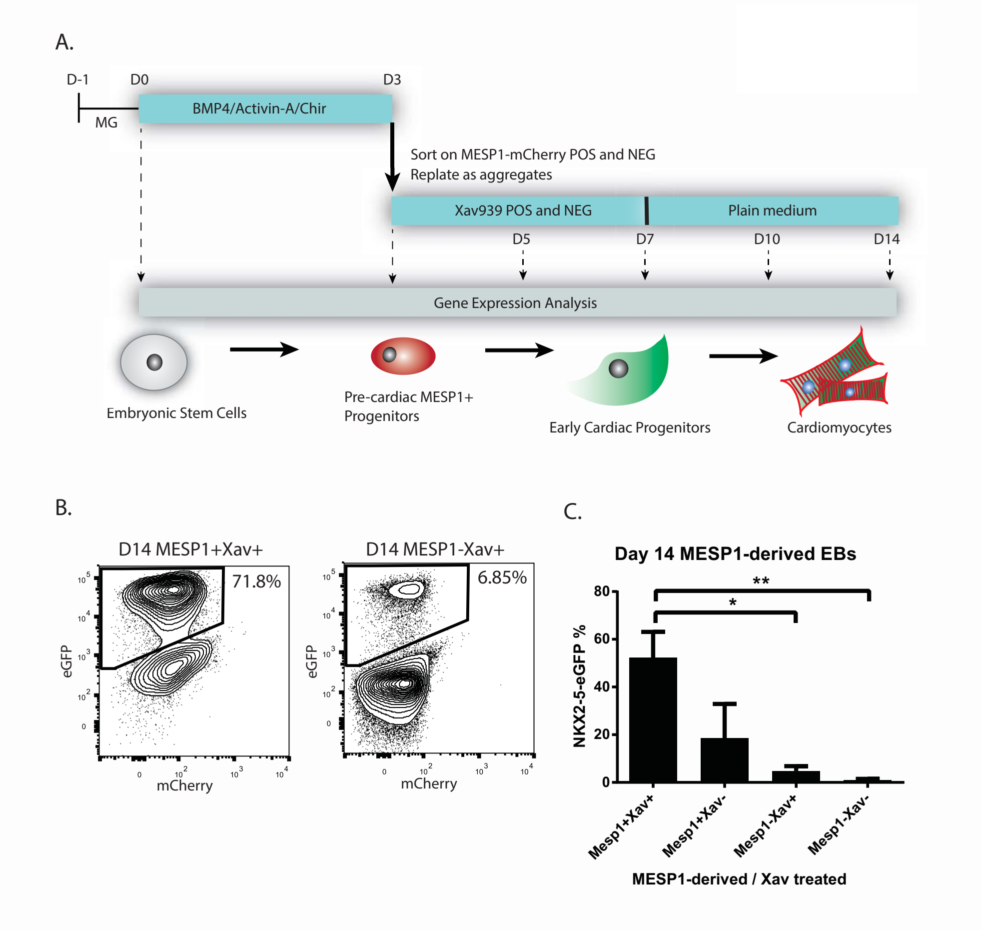
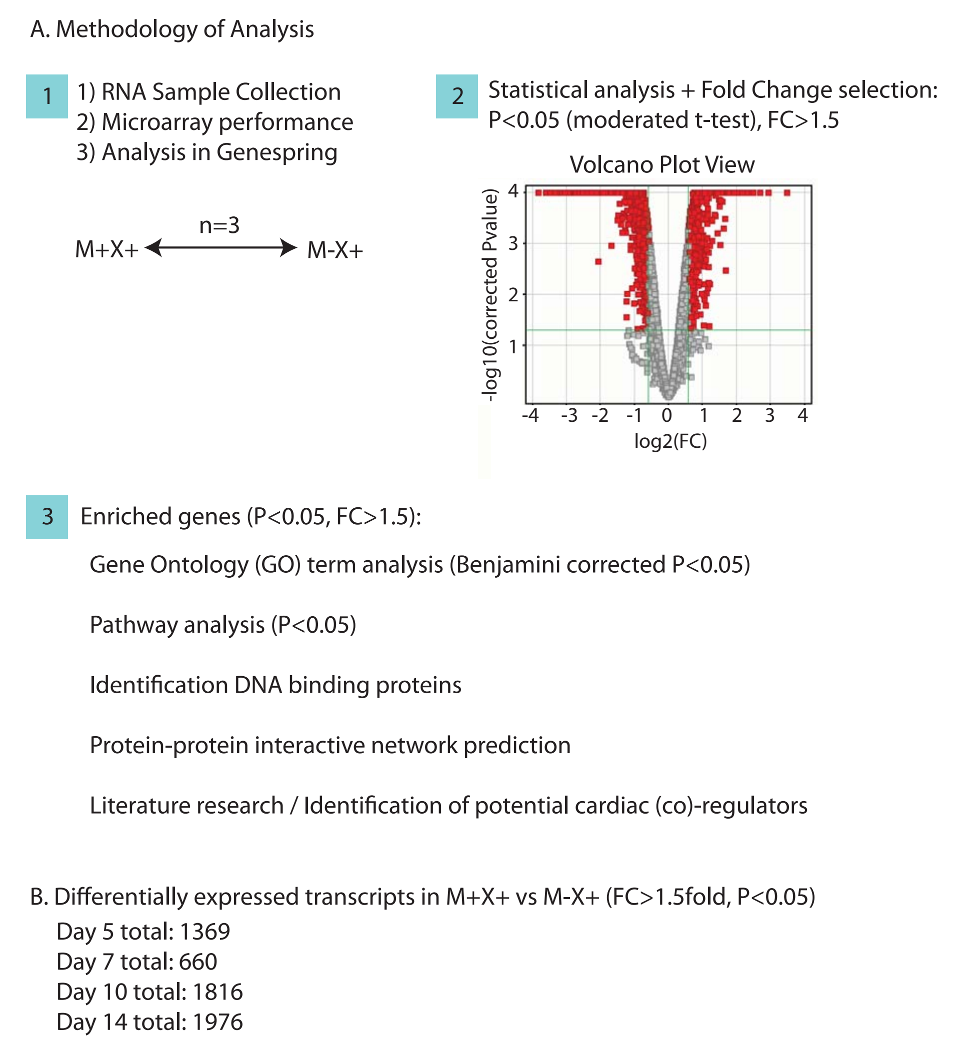


Figure 1. A: hESCs were differentiated as monolayer and were sorted at day 3 on their MESP1-mCherry expression. Positive and negative fractions were further differentiated and RNA was collected at sequential timepoints day 5, day 7, day 10, and day 14. **B,C:** The efficiency of the cardiac differentiations was monitored by flow cytometry at day 14 of differentiation. NKX2-5-eGFP levels were highly increased in the MESP1-positive derived fraction, treated with Wnt-pathway inhibitor Xav939. MESP1-negative fractions, lacking Xav939 treatment, were showing almost no NKX2-5-eGFP expression levels* : P<0.05, **: P<0.001. Error bars indicate SEM: standard error of the mean. N=3.


A. Methodology of Analysis

1 1) RNA Sample Collection
2) Microarray performance
3) Analysis in Genespring

M+X+ ← n=3 → M-X+

2 Statistical analysis + Fold Change selection:
P<0.05 (moderated t-test), FC>1.5

3 Enriched genes (P<0.05, FC>1.5):

Gene Ontology (GO) term analysis (Benjamini corrected P<0.05)

Pathway analysis (P<0.05)

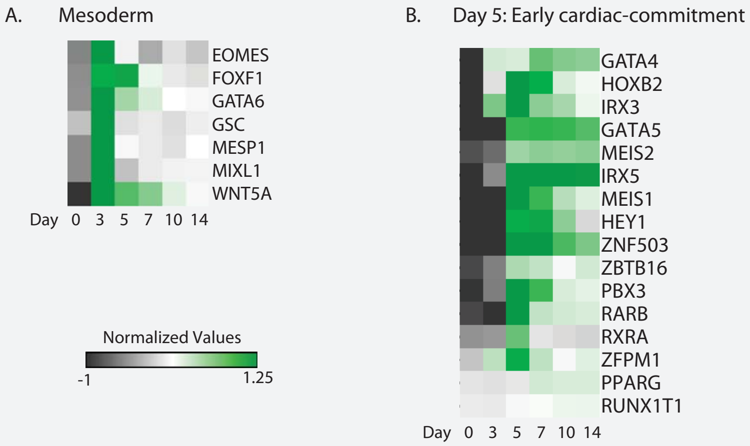
Identification DNA binding proteins

Protein-protein interactive network prediction

Literature research / Identification of potential cardiac (co)-regulators

B. Differentially expressed transcripts in M+X+ vs M-X+ (FC>1.5fold, P<0.05)
Day 5 total: 1369
Day 7 total: 660
Day 10 total: 1816
Day 14 total: 1976

Figure 2. A: Schematic overview of the methodology of gene selection from three biological replicate pairs, and the different methods of analysis that contribute to the identification of putative cardiac co-(regulators). A volcano plot view allows selection of differentially expressed genes, based on a P-value (P<0.05, moderated t-test), and a fold change > 1.5. **B**: The number of differentially expressed transcripts at each timepoint (M+X+ vs M-X+), based on a FC>1.5 fold difference, and a gene-specific p-value<0.05. N=3.



C. Protein-protein interaction network (STRING) - Enriched Transcription Factors Day 5

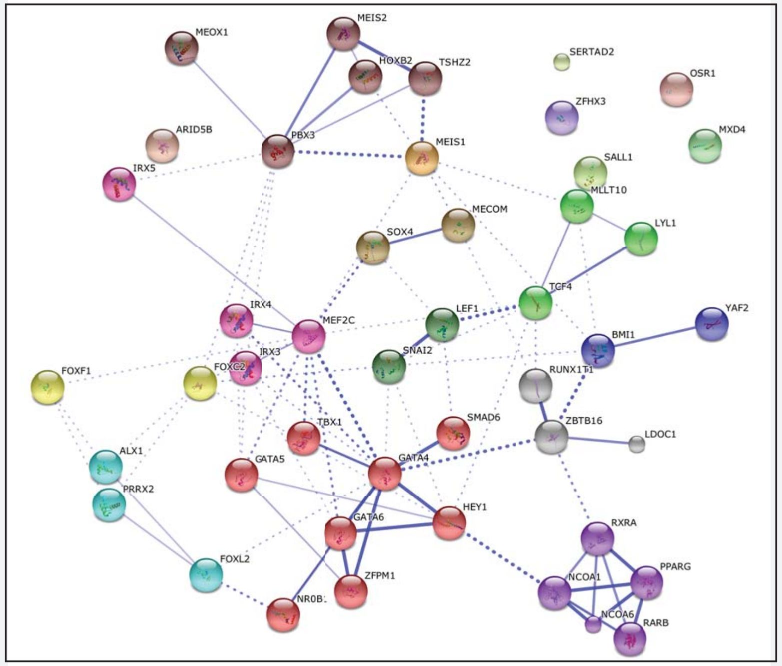
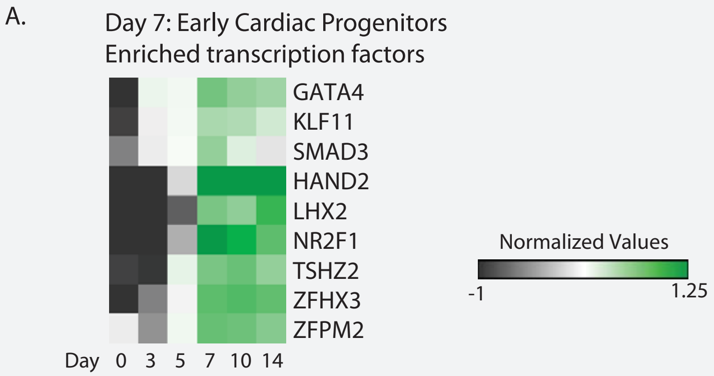


Figure 3. A. Heatmap visualization of the relative expression levels of mesoderm genes throughout cardiac differentiation, showing a stage-specific enrichment in MESP1-mCherry isolated progenitors at day 3 of differentiation. Heatmap shows averaged values from n=3. **B.** Relative expression levels of DNA binding transcriptional regulators that were enriched at day 5 of differentiation in the MESP1-mCherry positive derivatives. Genes were clustered based on a One Minus Pearson Correlation. Heatmap shows averaged values from n=3. **C.** Evidence for protein-protein interaction networks of enriched transcription factors at day 5 of differentiation was constructed by STRING. Interactions with a medium confidence > 0.4 are visualized. Proteins are clustered using the MCL algorithm. Every color represents a cluster. Inter-cluster edges are represented by dashed lines.



B. Protein-protein interaction network (STRING) - Enriched Transcription Factors Day 7

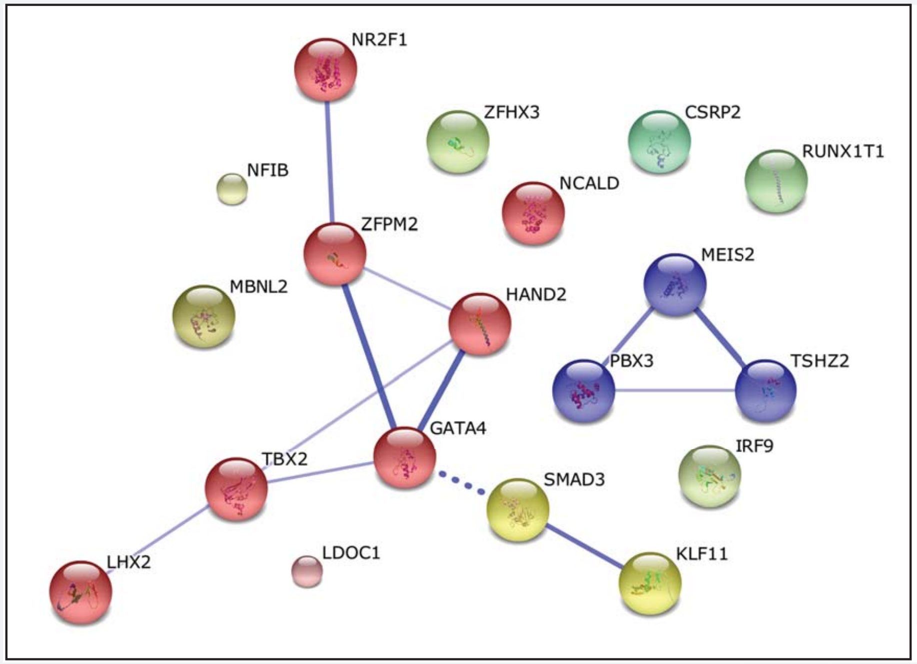


Figure 4. Enriched transcripts at day 7 of differentiation in MESP1-mCherry derivatives. Genes were selected based on a P-value> 0.05 and FC>1.5 fold difference in expression, when compared day 7 MESP1-mCherry negative derivatives. **A:** Enriched DNA binding transcripts at day 7 of differentiation in MESP1-mCherry derivatives. Heatmap shows averaged values from n=3. **B:** Evidence for protein-protein interaction networks of the enriched transcription factors at day 7 of differentiation was constructed by STRING. Interactions with a medium confidence > 0.4 are visualized. Proteins are clustered using the MCL algorithm. Every color represents a cluster. Inter-cluster edges are represented by dashed lines.

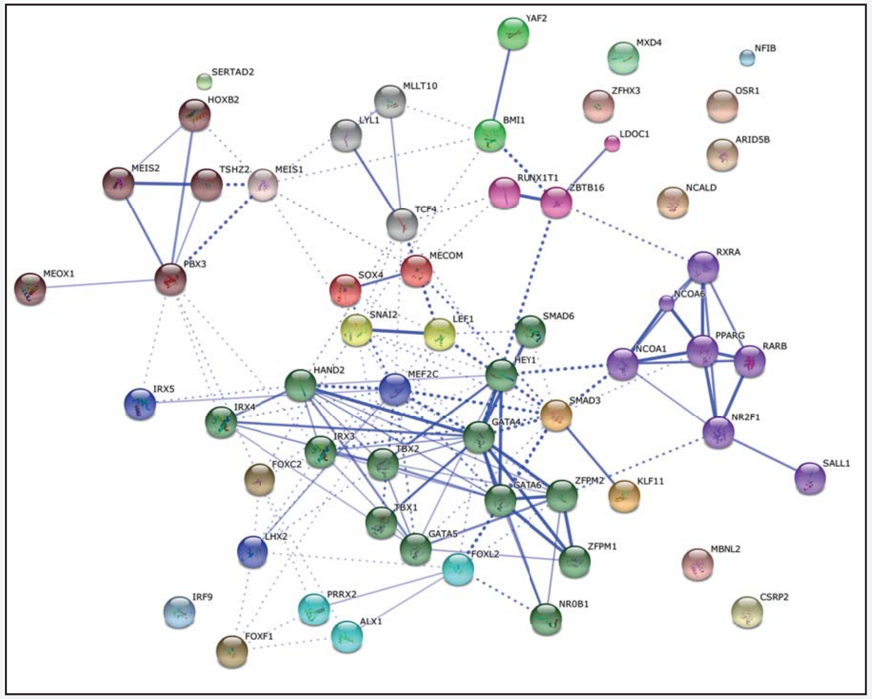


Figure 4C. Protein-protein interaction network (STRING) - Enriched Transcription Factors Day 5-7

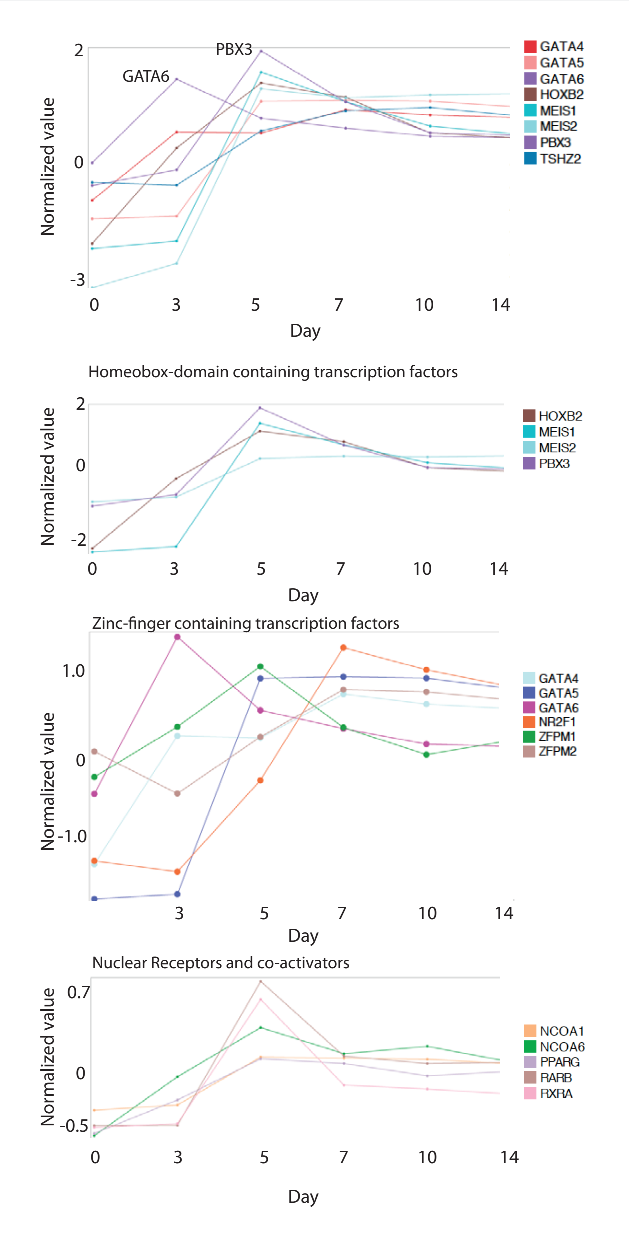


Figure 5. Profile Plots of Day 5 and Day 7 Enriched Transcription Factors. The upper panel shows the expression patterns upon cardiac differentiation of a selection of DNA binding transcription factors. The second upper panel shows the similarities between the expression patterns of HOXB2, MEIS1, MEIS2, and PBX3. The second last panel shows the expression pattners of GATA factors and ZFPM factors, and shows the similarities between ZFPM2 and GATA4 from day 7-14 of differentiation. The lower panel shows the expression levels of retinoic acid nuclear receptors and their co-activators, peaking at day 5 of differentiation. Normalized alues are visualized on a global expression scale and averaged from n=3 at each timepoint.

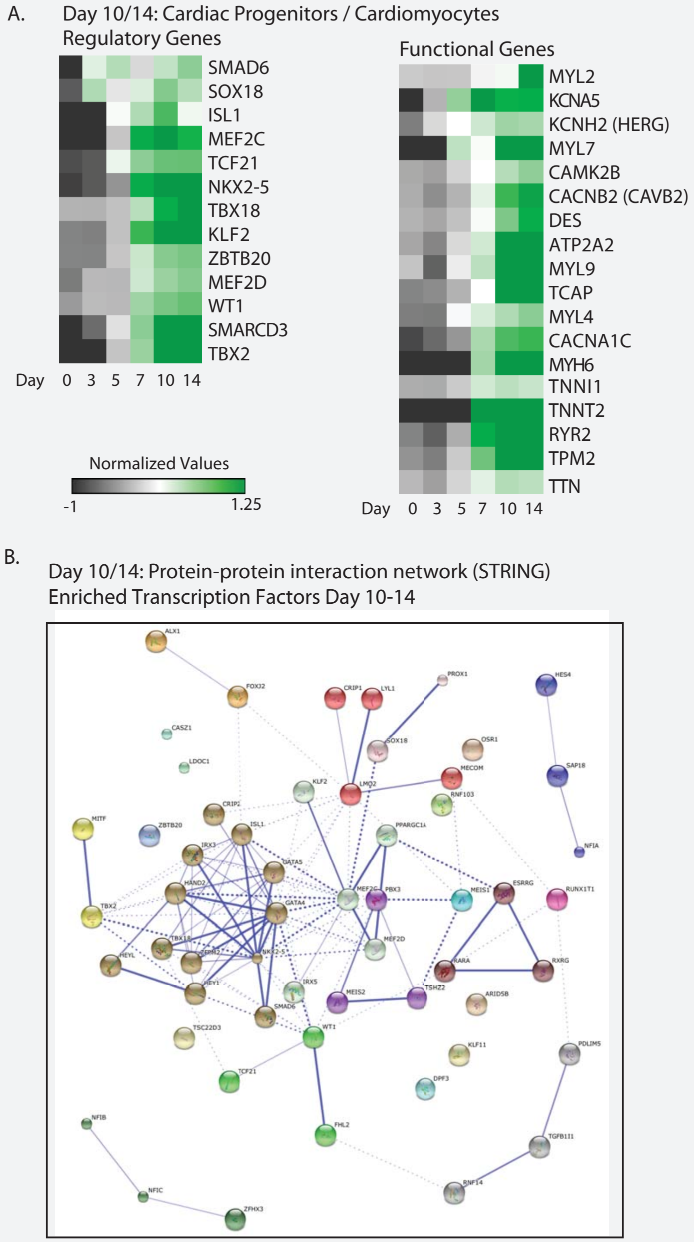


Figure 6. Enriched transcripts at day 10 and day 14 of differentiation in MESP1-mCherry derivatives. Genes were selected based on a P-value> 0.05 and a FC>1.5 fold difference in expression with day 10 and day 14 MESP1-mCherry negative derivatives. **A.** A selection of regulatory and functional genes that were enriched. Heatmaps show averaged values from n=3. **B.** Evidence for protein-protein interaction networks of the enriched transcription factors at day 10 and 14 of differentiation was constructed by STRING. Interactions with a medium confidence > 0.4 are visualized. Proteins are clustered using the MCL algorithm. Every color represents a cluster. Inter-cluster edges are represented by dashed lines.

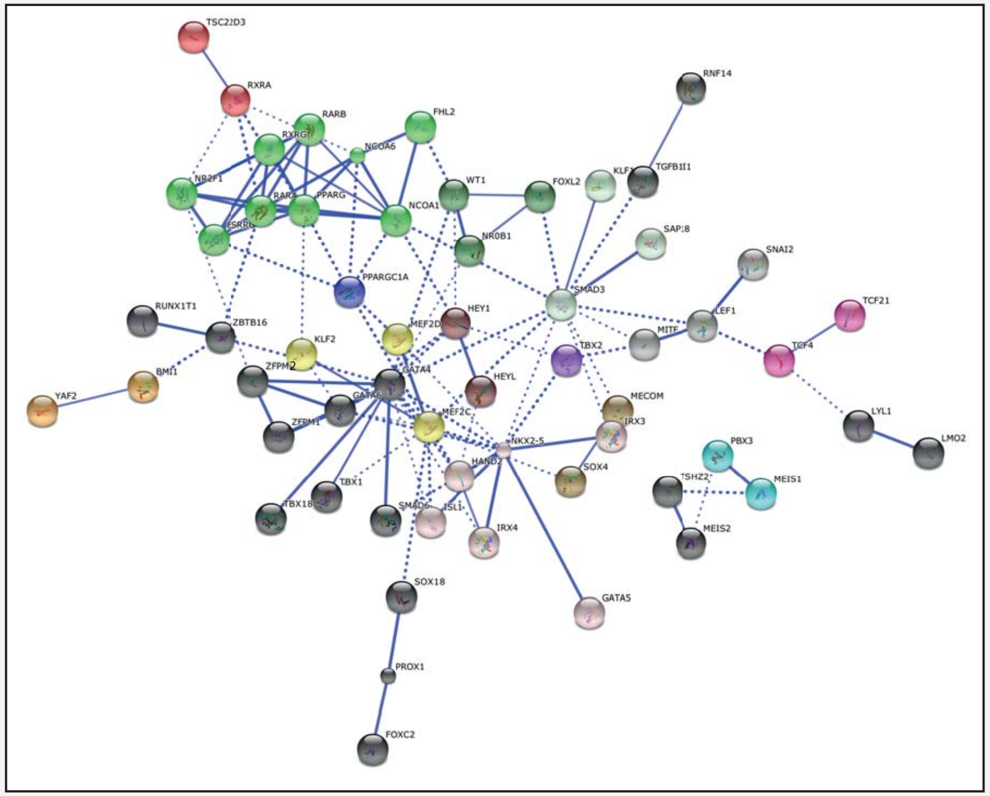
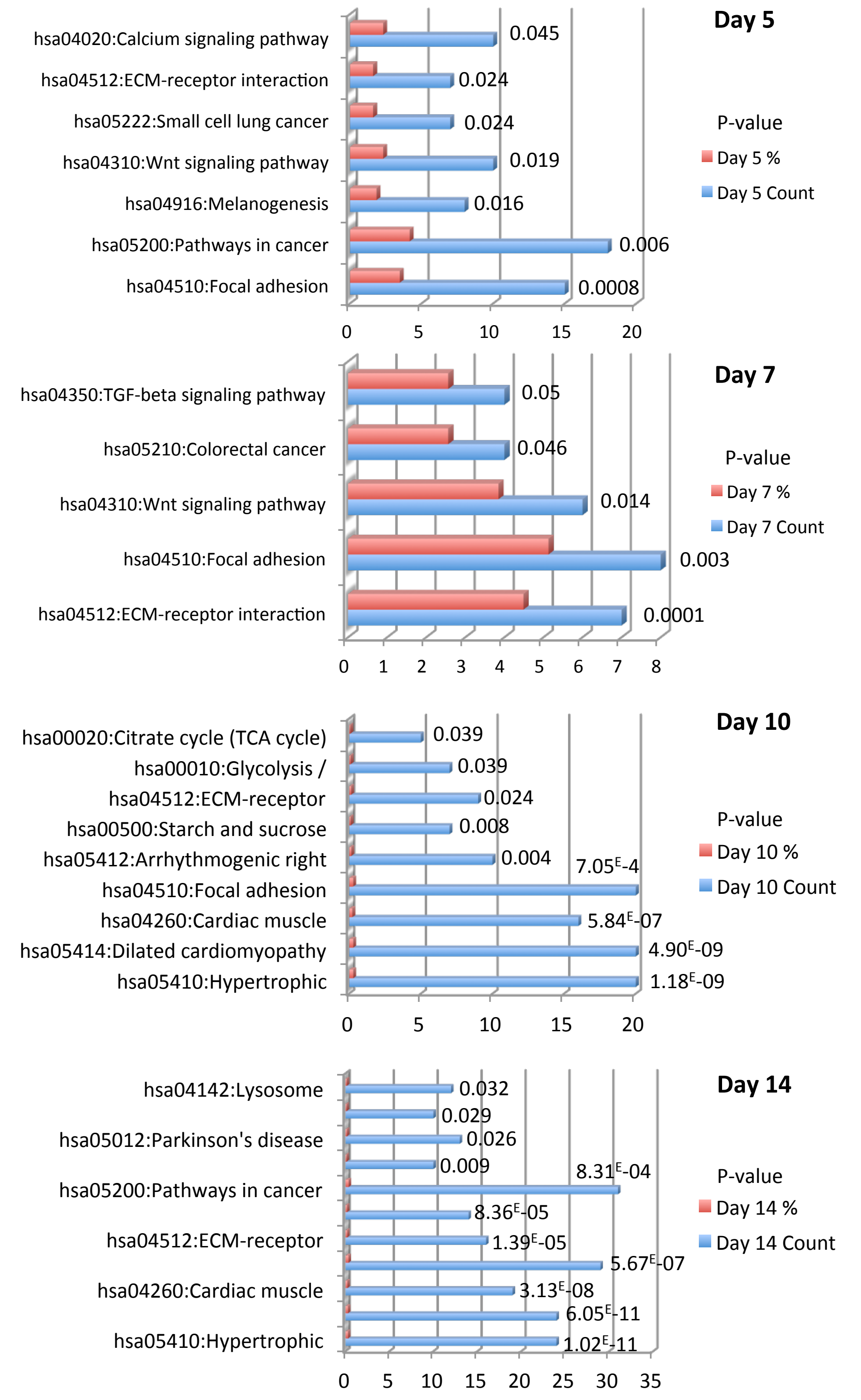


Figure 7. Merged protein-protein interaction network of the enriched transcription factors from all timepoints (d5, d7, d10, d14), obtained from STRING. Interactions with a high confidence > 0.7 are visualized. Proteins are clustered using the MCL algorithm. Every color represents a cluster. Inter-cluster edges are represented by dashed lines. Non-connected gene nodes are not visualized. We could identify the development of three distinct large node clusters, including retinoic acid nuclear receptors, containing PPARγ and RAR α/β, and two large interconnected cardiac networks centered on either GATA4 or NKX-5.



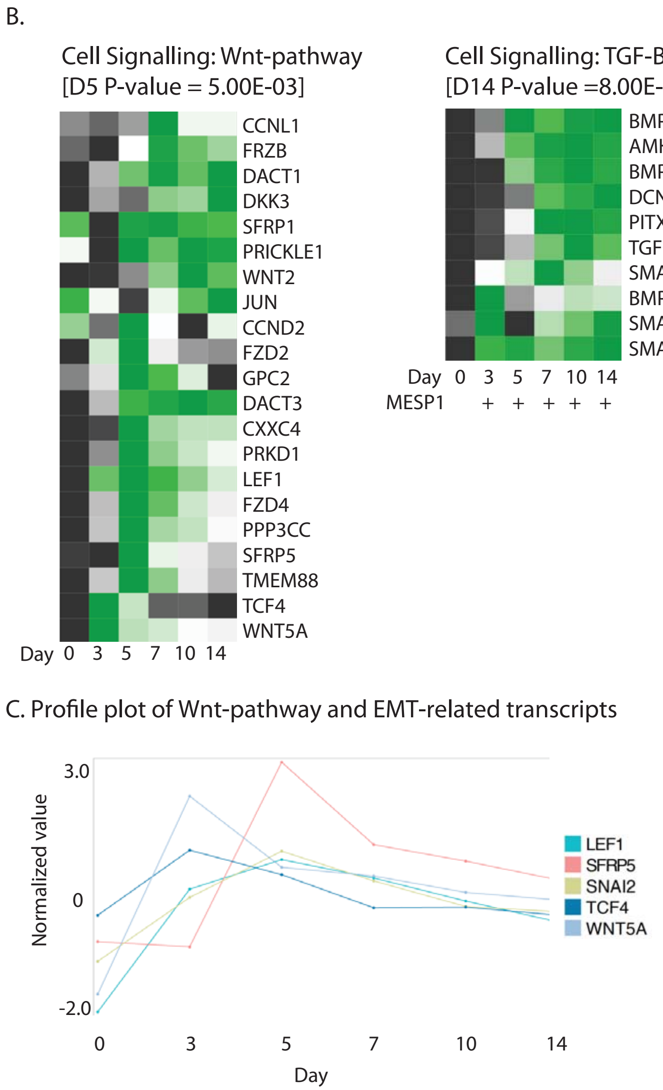


Figure 8. Enriched signalling pathways were identified through KEGG pathway analysis. A: Pathway analysis was performed on the enriched genes, selected as described before. Pathways that are visualized have a P<0.05. Gene count indicates the number of genes that could be identified for each enriched pathway. % indicates the percentile of the gene count from the total number of genes in the specific pathway. Numbers indicate the pathway-specific p-value. A comprehensive list of pathways can be found in Table S2. **B:** Heatmap visualization of the enriched transcripts belonging to the Wnt signalling pathway, which is enriched in early cardiac committed progenitors, and the TGF-β signalling pathway, enriched in cardiac progenitors and cardiomyocytes. Heatmaps show averaged values from n=3. **C:** Profile Plots of enriched Wnt-pathway-related and EMT-related gene expression levels throughout cardiac differentiation. Normalized values are visualized on a global expression scale and averaged on n=3.

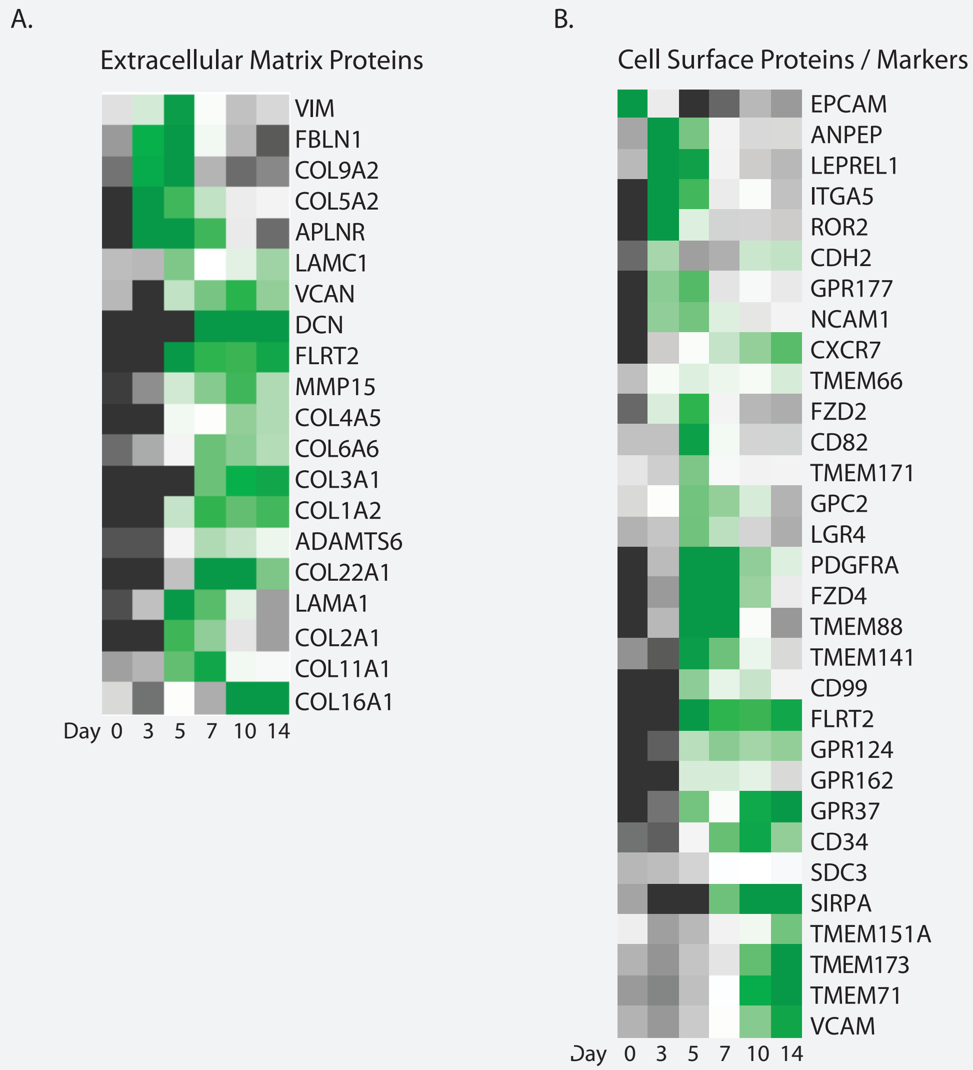


Figure 9. A: Heatmap visualizes the temporal enrichment of extracellular matrix proteins throughout cardiac differentiation. **B:** Visualization of temporal expressed cell membrane proteins to identify potential cell surface markers for subsets of cardiac progenitors or cardiomyocytes. Heatmaps show averaged values from n=3.

Day 5 (Enriched transcripts in M+ vs M-, FC>1.5, P<0.05)

Category	Term	Count	%	Benjamini Corrected P-value
PANTHER_BP	BP00193:Developmental processes	89	20,60	1,35E-06
PANTHER_BP	BP00102:Signal transduction	108	25,00	0,011389832
PANTHER_BP	BP00274:Cell communication	47	10,88	0,018208241
PANTHER_BP	BP00248:Mesoderm development	27	6,25	0,020433632
PANTHER_BP	BP00246:Ectoderm development	31	7,18	0,016956308
PANTHER_BP	BP00199:Neurogenesis	27	6,25	0,032211571
PANTHER_MF	MF00178:Extracellular matrix	27	6,25	3,52E-05
PANTHER_MF	MF00179:Extracellular matrix structural protein	11	2,55	0,001599702

Day 7 (Enriched transcripts in M+ vs M-, FC>1.5, P<0.05)

Category	Term	Count	%	Benjamini Corrected P-value
PANTHER_BP	BP00193:Developmental processes	38	24,36	0,002291487
PANTHER_BP	BP00124:Cell adhesion	16	10,26	0,015033261
PANTHER_BP	BP00281:Oncogenesis	12	7,69	0,059284076
PANTHER_BP	BP00274:Cell communication	22	14,10	0,046309515
PANTHER_BP	BP00102:Signal transduction	43	27,56	0,187888803
PANTHER_BP	BP00122:Ligand-mediated signaling	9	5,77	0,425197991
PANTHER_BP	BP00071:Proteolysis	15	9,62	0,405883178
PANTHER_BP	BP00109:Receptor protein serine/threonine kinase signaling pathwa	3	1,92	0,531826987
PANTHER_BP	BP00248:Mesoderm development	10	6,41	0,535902255
PANTHER_BP	BP00103:Cell surface receptor mediated signal transduction	21	13,46	0,653573879
PANTHER_BP	BP00267:Homeostasis	5	3,21	0,667475861
PANTHER_MF	MF00178:Extracellular matrix	18	11,54	1,59E-06
PANTHER_MF	MF00282:Aspartic protease	4	2,56	0,035633978
PANTHER_MF	MF00179:Extracellular matrix structural protein	6	3,85	0,041276498

Table 1B. Gene Ontology Analysis of enriched transcripts at sequential stages of differentiation (P<0.05).

Day 10 (Enriched transcripts in M+ vs M-, FC>1.5, P<0.05)

Category	Term	Count	%	Benjamini Corrected P-value
PANTHER_BP	BP00173:Muscle contraction	38	5,65	2,02E-14
PANTHER_BP	BP00250:Muscle development	25	3,71	9,55E-08
PANTHER_BP	BP00248:Mesoderm development	51	7,58	1,36E-06
PANTHER_BP	BP00193:Developmental processes	122	18,13	1,58E-04
PANTHER_BP	BP00251:Heart development	10	1,49	0,004440264
PANTHER_BP	BP00285:Cell structure and motility	65	9,66	0,019756301
PANTHER_BP_	BP00287:Cell motility	27	4,01	0,018914324
PANTHER_MF	MF00261:Actin binding cytoskeletal protein	48	7,13	1,83E-11
PANTHER_MF	MF00091:Cytoskeletal protein	63	9,36	6,27E-06
PANTHER_MF	MF00178:Extracellular matrix	34	5,05	2,34E-04
PANTHER_MF	MF00230:Actin binding motor protein	9	1,34	0,024317475

Table 1. Gene Ontology Analysis of enriched transcripts at sequential stages of differentiation. Differentially expressed genes in the MESP1-mCherry positive derived cardiac populations were selected on a FC>1.5 fold difference with their MESP1-mCherry negative counterpart, and a P-value<0.05, n=3. GO-analysis was performed using the Panther Classification System. Multiple-GO-term correction was performed using the Benjamini-Yekutieli method. GO-terms hold a corrected p-value <0.05.

REFERENCES

1. Beqqali, A., Kloots, J., Ward-van Oostwaard, D., Mummery, C. & Passier, R. Genome-Wide Transcriptional Profiling of Human Embryonic Stem Cells Differentiating to Cardiomyocytes. STEM CELLS **24**, 1956–1967 (2006).
2. Paige, S. L. et al. A Temporal Chromatin Signature in Human Embryonic Stem Cells Identifies Regulators of Cardiac Development. Cell **151**, 221–232 (2012).
3. Elliott, D. A. et al. NKX2-5eGFP/w hESCs for isolation of human cardiac progenitors and cardiomyocytes. Nat Meth **8**, 1037–1040 (2011).
4. Hartogh, Den, S. C. et al. Dual Reporter MESP1 mCherry/w-NKX2-5 eGFP/whESCs Enable Studying Early Human Cardiac Differentiation. STEM CELLS **33**, 56–67 (2014).
5. Curtis, R. K., Oresic, M. & Vidal-Puig, A. Pathways to the analysis of microarray data. Trends Biotechnol. 23, 429–435 (2005).
6. Szklarczyk, D. et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Research **43**, D447–52 (2015).
7. Jacobs, Y., Schnabel, C. A. & Cleary, M. L. Trimeric association of Hox and TALE homeodomain proteins mediates Hoxb2 hindbrain enhancer activity. Mol. Cell. Biol. **19**, 5134–5142 (1999).
8. Chang, C. P., Brocchieri, L. & Shen, W. F. Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus. … and cellular biology (1996).
9. Phelan, M. L., Rambaldi, I. & Featherstone, M. S. Cooperative interactions between HOX and PBX proteins mediated by a conserved peptide motif. Mol. Cell. Biol**.** **15**, 3989–3997 (1995).
10. Stankunas, K. et al. Pbx/Meis deficiencies demonstrate multigenetic origins of congenital heart disease. Circulation Research **103**, 702–709 (2008).
11. Laforest, B., Bertrand, N. & Zaffran, S. Anterior Hox Genes in Cardiac Development and Great Artery Patterning. JCDD **1**, 3–13 (2014).
12. Maves, L. et al. Pbx homeodomain proteins direct Myod activity to promote fast-muscle differentiation. Development **134**, 3371–3382 (2007).
13. Berkes, C. A. et al. Pbx marks genes for activation by MyoD indicating a role for a homeodomain protein in establishing myogenic potential. Mol. Cell **14**, 465–477 (2004).
14. Lindsley, R. C., Gill, J. G., Murphy, T. L., Langer, E. M. & Cai, M. Mesp1 coordinately regulates cardiovascular fate restriction and epithelial-mesenchymal transition in differentiating ESCs. Cell Stem Cell **3** 55–68 (2008).
15. Cantor, A. B. & Orkin, S. H. Coregulation of GATA factors by the Friend of GATA (FOG) family of multitype zinc finger proteins. Semin. Cell Dev. Biol. **16**, 117–128 (2005).
16. Walton, R. Z., Bruce, A., Olivey, H. E. & Najib, K. Fog1 is required for cardiac looping in zebrafish. Developmental **289**, 482–493 (2006).
17. Melnick, A. et al. AML-1/ETO fusion protein is a dominant negative inhibitor of transcriptional repression by the promyelocytic leukemia zinc finger protein. Blood **96**, 3939–3947 (2000).
18. Zhang, L. et al. Characterization of a t(5;8)(q31;q21) translocation in a patient with mental retardation and congenital heart disease: implications for involvement of RUNX1T1 in human brain and heart development. Eur. J. Hum. Genet. 17, 1010–1018 (2009).
19. Huggins, G. S., Bacani, C. J., Boltax, J., Aikawa, R. & Leiden, J. M. Friend of GATA 2 physically interacts with chicken ovalbumin upstream promoter-TF2 (COUP-TF2) and COUP-TF3 and represses COUP-TF2-dependent activation of the atrial natriuretic factor promoter. J. Biol. Chem. 276, 28029–28036 (2001).
20. Lu, J. R. et al. FOG-2, a heart- and brain-enriched cofactor for GATA transcription factors. Mol. Cell. Biol. 19, 4495–4502 (1999).
21. Klinedinst, S. L. & Bodmer, R. Gata factor Pannier is required to establish competence for heart progenitor formation. Development 130, 3027–3038 (2003).
22. Berry, F. B., Miura, Y., Mihara, K., Kaspar, P. & Sakata, N. Positive and negative regulation of myogenic differentiation of C2C12 cells by isoforms of the multiple homeodomain zinc finger transcription factor ATBF1. Journal of Biological … 25057–25065 (2001).
23. Liu, Y. et al. Genetic polymorphisms in ZFHX3 are associated with atrial fibrillation in a Chinese Han population. PLoS ONE 9, e101318 (2014).
24. Benjamin, E. J. et al. Variants in ZFHX3 are associated with atrial fibrillation in individuals of European ancestry. Nat Genet 41, 879–881 (2009).
25. 2Martin, R. I. R. et al. Chromosome 16q22 variants in a region associated with cardiovascular phenotypes correlate with ZFHX3 expression in a transcript-specific manner. BMC Genet. 15, 136 (2014).
26. Lian, X. et al. Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling(2). Proceedings of the … 109, 1848–1857 (2012).
27. Witty, A. D. et al. Generation of the epicardial lineage from human pluripotent stem cells. Nature Biotechnology 32, 1026–1035 (2014).
28. Gessert, S. & Kühl, M. The multiple phases and faces of wnt signaling during cardiac differentiation and development. Circulation Research 107, 186–199 (2010).
29. 29. beta in cardiovascular development and function. Cytokine Growth Factor Rev. 14, 391–407 (2003).
30. Medici, D., Hay, E. D. & Goodenough, D. A. Cooperation between snail and LEF-1 transcription factors is essential for TGF-beta1-induced epithelial-mesenchymal transition. Mol. Biol. Cell 17, 1871–1879 (2006).
31. Cheng, P. et al. Fibronectin mediates mesendodermal cell fate decisions. Development 140, 2587–2596 (2013).
32. Lockhart, M., Wirrig, E., Phelps, A. & Wessels, A. Extracellular matrix and heart development. Birth Defects Research Part A: Clinical and Molecular Teratology 91, 535–550 (2011).
33. Farouz, Y., Chen, Y., Terzic, A. & Menasché, P. Concise review: growing hearts in the right place: on the design of biomimetic materials for cardiac stem cell differentiation. STEM CELLS 33, 1021–1035 (2015).
34. Turbendian, H. K. et al. GATA factors efficiently direct cardiac fate from embryonic stem cells. Development 140, 1639–1644 (2013).
35. Palpant, N. J., Pabon, L., Rabinowitz, J. S., Hadland, B. K., Stoick-Cooper, C. L., Paige, S. L., et al. (2013). Transmembrane protein 88: a Wnt regulatory protein that specifies cardiomyocyte development. Development, 140, 3799–3808.
36. Schoore, G. V., Mendive, F., Pochet, R. & Vassart, G. Expression pattern of the orphan receptor LGR4/GPR48 gene in the mouse. Histochem Cell Biol 124, 35–50 (2005).
37. Loh, E. D., Broussard, S. R. & Kolakowski, L. F. Molecular characterization of a novel glycoprotein hormone G-protein-coupled receptor. Biochemical and Biophysical Research Communications 282, 757–764 (2001).
38. Evseenko, D. et al. Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. Proceedings of the National Academy of Sciences 107, 13742–13747 (2010).
39. Paxton, C. N., Bleyl, S. B., Chapman, S. C. & Schoenwolf, G. C. Identification of differentially expressed genes in early inner ear development. Gene Expr. Patterns 10, 31–43 (2010).
40. 40. Descamps, B. et al. Frizzled 4 regulates arterial network organization through noncanonical Wnt/

planar cell polarity signaling. Circulation Research 110, 47–58 (2012).

41. 41. Müller, P.-S. et al. The fibronectin leucine-rich repeat transmembrane protein Flrt2 is required in the epicardium to promote heart morphogenesis. Development 138, 1297–1308 (2011).

42. 42. Miller, R. A., Christoforou, N., Pevsner, J., McCallion, A. S., & Gearhart, J. D. (2008). Efficient Array-Based Identification of Novel Cardiac Genes through Differentiation of Mouse ESCs. PLoS ONE, 3(5), e2176.

43. 43. Palty, R., Raveh, A., Kaminsky, I., Meller, R., & Reuveny, E. (2012). SARAF Inactivates the Store Operated Calcium Entry Machinery to Prevent Excess Calcium Refilling. Cell, 149(2), 425–438. (2012).

44. 44. Dubois, N. C. et al. SIRPA is a specific cell-surface marker for isolating cardiomyocytes derived from human pluripotent stem cells. Nature Biotechnology 29, 1011–1018 (2011).

45. 45. Skelton, R., Costa, M., Anderson, D. J., & Bruveris, F. (2014). SIRPA, VCAM1 and CD34 identify discrete lineages during early human cardiovascular development. Stem Cell Research, 13, 172–179.

46. 46. Amigo, J. D. et al. The role and regulation of friend of GATA-1 (FOG-1) during blood development in the zebrafish. Blood 114, 4654–4663 (2009).

47. 47. Orkin, S. H. Diversification of haematopoietic stem cells to specific lineages. Nat Rev Genet 1, 57–64 (2000).

48. 48. Chan, S. S.-K. et al. Mesp1 Patterns Mesoderm into Cardiac, Hematopoietic, or Skeletal Myogenic Progenitors in a Context-Dependent Manner. Cell Stem Cell 12, 587–601 (2013).

49. 49. Saga, Y. Mesp1 Expression Is the Earliest Sign of Cardiovascular Development. TCM 10, 345–352 (2001).

50. 50. Carmon, K. S., Gong, X., Lin, Q., Thomas, A. & Liu, Q. R-spondins function as ligands of the orphan receptors LGR4 and LGR5 to regulate Wnt/β-catenin signaling. PNAS 108, 11452–11457 (2011).

51. 51. Yi, J. et al. Analysis of LGR4 receptor distribution in human and mouse tissues. PLoS ONE 8, e78144 (2013).

52. 52. Barker, N., Tan, S. & Clevers, H. Lgr proteins in epithelial stem cell biology. Development 140, 2484–2494 (2013).

53. 53. Brohée, S. & van Helden, J. Evaluation of clustering algorithms for protein-protein interaction networks. BMC Bioinformatics 7, 488 (2006).

54. 54. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504 (2003).

CHAPTER 5

Culturing Cardiac Conduction System Cells from Human Pluripotent Stem Cells using TBX3 regulatory elements: Technical and Biological Challenges

Sabine den Hartogh[1], Tom Zwetsloot[1], Henk van Weerd[2],
Vincent Christoffels[2], Robert Passier[1,3]

1. Leiden University Medical Centre, Dept. Anatomy and Embryology, Leiden, The Netherlands.
2. Amsterdam Medical Centre, Heart Failure Research Centre, Amsterdam, The Netherlands
3. Department of Applied Stem cell Technologies. MIRA Institute for Biomedical Technology and Technical Medicine. University of Twente, P.O.Box 217, Enschede, The Netherlands.

ABSTRACT

Human pluripotent stem cell (hPSC)-derived cardiovascular cells hold great expectations for the field of regenerative therapy of the heart, disease modelling, drug development, and toxicity screenings. Recent optimized cardiomyocyte differentiation protocols yield high percentage of either atrial or ventricular cardiomyocytes, whereas only a small subset of cardiomyocytes poses a pacemaker-like identity. Although studies indicate a role for Neuregulin/ErB and Notch signalling in the formation of specialized conduction-like cells, detailed insights into the regulation of cardiac conduction cell formation is largely unknown. In the mammalian heart, TBX3 is expressed specifically in cells from the cardiac conduction system (CCS), which initiate and propagate the electric impulse required for the coordinated contraction of the chambers. Previously it was shown that two enhancer elements are responsible for atrioventricular TBX3 expression *in vivo.* Here, we incorporated these enhancer elements in a reporter construct, coupled to a minimal promoter, and driving mCherry expression. Transgenic lines were generated in the background of the previously described NKX2-5-eGFP reporter hESC line. Subsequently, we screened this double reporter line with cytokines/molecules in order to identify factors/pathways that may be important for activating TBX3-positive pacemaker-like cardiomyocytes.

Introduction

The heart is comprised of cardiomyocytes, which can be subdivided in two classes: working myocardium (atrial and ventricular), and conductive tissue. In higher vertebrates the conduction system consists of a central and a peripheral component. In the central conduction system (CCS), the sinoatrial node (SAN) initiates electrical impulses, that are propagated through the atria, and delayed in the atrioventricular node (AVN). This is followed by rapid propagation through the interventricular ring, the right atrioventricular ring bundle, the atrioventricular bundle, and the proximal bundle branches (BBs). The impulses then enter the peripheral conduction: the distal part of the BBs and the peripheral ventricular conduction network. Cardiomyocytes of the conduction system are highly specialized cells. Their development is tightly controlled by specific regulatory pathways [1]. Several transcription factors play crucial roles in the differentiation of cells of the CCS, including T-box transcription factors. Alterations in transcriptional networks during formation of the conduction system cause electrical disturbances, such as sinus node dysfunction or atrioventricular block[2-4] may lead to life-threatening cardiac arrhythmias. Insights into the molecular control of the developing conduction system will give us more understanding into the onset of congenital and acquired conductive abnormalities. The T-box factor TBX3 is specifically expressed in the developing and mature CCS in mammals. During development, TBX3 is expressed in the cardiac inflow and outflow tract, the SAN, the AVN, internodal regions, atrioventricular regions, the AV bundle, and in the BBs[5]. TBX3 is a repressor of working myocardial genes, such as NPPA (ANF) and Connexin40 (CX40), and is described as critical factor in the formation of the SAN by repressing the atrial working myocardial gene expression, and by activating SAN genes, such as LBH, HCN family members, and Connexin 30.2[6]. Horsthuis et al. showed how a transgenic mouse reporter line, harbouring a 160Kb-BAC sequence spanning the Tbx3 locus, displayed gene expression only in a subdomain of TBX3 positive cardiac cells, including the primordial AVN, ventral/cranial, and right side of the AVC. This implies that the activity of specific regulatory DNA elements are responsible for a precise pattern of TBX3 expression in the CCS[7]. In a recent study by the same group, incorporation of upstream sequences of this TBX3 locus complementary expression was found in the complete AV rings and in the prospective AV bundle[8]. They identified two enhancer regions eA and eB to be responsible for the AVN-TBX3 expression, which were activated by both BMP signalling and GATA4 expression.

For therapeutic and pharmaceutical applications, including cardiac repair and/or drug development/discovery, it will be of interest to derive cells from the cardiac conduction system *in vitro*. Therefore, in this project, we studied the *in vitro* derivation of CCS cells from pluripotent stem cells, by generating a TBX3eAeB-Hsp68 (TBX3e)-mCherry reporter in the previously generated NKX2-5-eGFP-hESC line[9]. Here, both enhancer regions were sequentially coupled to promoter-element Hsp68 and sequences encoding for fluorescent reporter protein mCherry. By screening the activation and inhibition of different signalling pathways at specific timings during cardiac differentiation, we investigated activation of TBX3e-mCherry expressing cells.

MATERIAL AND METHODS

Generation of Lentiviral TBX3eAeB-mCherry Reporter Construct

The TBX3e-mCherry reporter construct was generated through the Gateway® Cloning Technology, using the Gateway® LR Clonase® II Enzyme mix (Life Technologies). Two pENTR vectors were kindly provided by H. Van Weerd: pENTR4-mCherry-polyA and pENTR5-TBX3eAeB-Hsp68 (sequences can be found in **Supplemental Fig. 1**), which were used to clone mCherry-polyA and TBX3eAeB-Hsp68 into destination vector pLenti6-R4R2-V5-DEST **(Fig. 1a)**.

Enzyme mix:

	pENTR4 – mCherry	100ng	1.5uL
	pENTR5- eAeB	100ng	1.5uL
	pLenti6/R4R2/V5-DEST	150ng	1uL
	TE PH 8.0		4uL
	Total		8uL
+	**Enzyme mix**		**2uL**

The reaction was incubated at RT for 18 hours, and was stopped by addition of 1uL Proteinase K. A reaction volume of 2 uL was transformed into E. Coli (One Shot® TOP10 Chemically Competent E. coli, Life Technologies). Positive bacterial colonies were selected on Ampicillin-containing agar plates (100ug/mL). Destination vectors were isolated from selected colonies and digested with EcoRV (New England Biolabs). Clones # 1, 2, 5, 7, 9, 10, 11, 13, 14, 16, and 17 were showing the expected bands (**Fig. 1b**). This was confirmed by a second digestion with PvuI (New England Biolabs) on clones # 7, 9, 10, 11, 13, 14, 16, and 17 (**Fig. 1c**). Clones # 7, 9, and 10 were sent for sequencing, from which clone 7 and 9 contained correct sequences. Clone 9 was used for further experiments. Lentiviral particles were produced by co-transfection with 3rd generation lentiviral packaging vectors (pMD2.G, pRRE and pRSV/REV) into HEK293T cells (Thermo Scientific). For transfections, the construct DNA was mixed with Lipofectamine 2000 in OptiMEM (Life Technologies) and was added to the media (DMEM containing 4% FCS). After 18 hours, cells were refreshed with DMEM-4% FCS. Lentiviruses were harvested 72 h after transfection and concentrated in a Beckman coulter ultracentrifuge at 20,000 rpm and at 40C for 2 hours. Lentiviruses were resuspended in 1% BSA-PBS and stored at -80°C. The titer was determined with a qPCR lentiviral titration kit (ABM Inc.).

Lentiviral transfection

To generate the TBX3e-mCherry reporter line, DeltaN3-NKX2-5$^{eGFP/w}$ hESCs were transduced with the lentiviral construct containing the TBX3e^{mCherry} expression cassette (**Fig. 2a**). The viral titer was estimated by qPCR on 4.2x106 IU/uL (ABM Inc.). We added 0 MOI (control), 10 MOI (0.25 uL), 20 MOI (0.5uL), and 50 MOI (1.2uL) to 100k hESCs. Incubation was performed in an Eppendorf tube for 5 minutes at RT, and after that, cells were plated in a 12 wells plate, coated with MEFs. After transduction, single clones were selected using blasticidin for 48 hours (Invitrogen, 10ug/mL). Clones were screened for the presence of mCherry protein by PCR. As positive control, we used a mCherry.pcDNA3.1 plasmid. PCR screening sequences were PCR_F1: GCATGGACGAGCTGTACAAG, and PCR_R1: GCATGAACTCCTTGATGATG. From the 32 screened colonies, we found 12 colonies positive, which was confirmed by a repeating PCR (**Fig. 2b, c**).

Culture and selection of positive clones

After blasticidin selection and PCR screening, multiple clones were kept as colony growing cultures. The cells were cultured on mouse embryonic fibroblasts (MEFs) in 2.85 cm2 organ dishes (Falcon), in 1 mL hESC medium (DMEM F12, Non-Essential Amino Acids, Knock Out Serum (Gibco by Life Technologies); bFGF (Miltenyi Biotec)) and were mechanically passaged once a week. Due to the random integration properties of lentiviral particles, three clones (A4, A5, and B5) were differentiated to cardiac cells by a standard cardiac differentiation protocol as described below. At day 14 of differentiation, NKX2-5-eGFP and TBX3e-mCherry expression was analysed by flow cytometry.

Cardiac Differentiation Protocol

One day prior to the differentiation, hESCs were passaged onto a 12-well tissue culture plate (Greiner CELLSTAR), coated with Matrigel (BD biosciences; growth factor reduced, phenol red-free), in 1 mL of hESC medium per well. The first day of differentiation, day 0, hESC medium was replaced by low insulin, serum–free medium BPEL (BSA, polyvinyl alcohol, essential lipids, as previously described[9, 10], containing the cytokines BMP4 (30 ng/mL, R&D Systems), Activin A (30 ng/mL, Miltenyi Biotec) and CHIR99021 (1.5µM, AxonMedchem). At day 3 of differentiation, the medium was refreshed with BPEL containing XAV939 (5µM, R&D Systems). At day 7, cells were refreshed with plain BPEL and further differentiated. In order to direct the differentiation towards the formation of CCS cells, this protocol was altered by the addition of several other cytokines in a range of concentrations at different timings as indicated. Cytokines that were screened included NRG-1β/ErbB signalling inhibitor AG-1478 (TEBU), FGFR antagonist SU5402 (Sigma-Aldrich Chemie) and WNT11 (R&D).

Flow Cytometry

To screen for the formation of TBX3e-mCherry expressing cells in the different differentiation protocols, fluorescent expression was analysed by flow cytometry at day 14 of differentiation. For this, the cells were enzymatically dissociated using 1x TrypLE Select (Life Technologies). After dissociation, cells were resuspended in FACS sorting buffer (PBS with 0.5% BSA and 2.5 mM EDTA) and analysed using a MACS Quant VYB flow cytometer (Miltenyi Biotec) with a 488 nm and a 561 nm laser. FACS data was further processed using FlowJo Software v10.

RESULTS

Generation of a Cardiac Conduction Reporter Line

In order to identify cells from the CCS in current cardiac differentiation protocols, and to develop protocols specifically directing towards cardiac conduction system lineages, we developed a cardiac conduction reporter construct. Enhancers eA and eB, recently identified to regulate TBX3 expression in the developing AV node during mouse embryogenesis, were both sequentially coupled to minimal promoter-element Hsp68 and sequences encoding for fluorescent reporter protein mCherry. The previously generated NKX2-5-eGFP hESCs[9] were transduced with lentiviral particles containing this reporter construct (**Fig. 2a**), resulting in a dual cardiac reporter line (**Fig. 2b**).

TBX3e-mCherry activation during In vitro cardiac differentiation

After blasticidin selection, 32 clones were picked and screened by PCR for the insertion of the mCherry cassette, of which 12 clones were positive (**Fig. 2c**). Due to the random integration properties of lentiviral particles and risk of transgene silencing, three clones (A4, A5, and B5) were further used for screening on TBX3e-mCherry expression upon differentiation.

As eA and eB were shown to be activated by both BMP signalling and GATA4 expression during mouse development[8], we supplemented our standard cardiac differentiation protocol with 20 ng/mL BMP4 from day 3 up to day 5, directly after mesoderm formation, hypothesising that cardiac conduction progenitors segregate from this early population, and taking into account endogenous GATA4 expression during differentiation[11, 12] (**Fig. 3a**). In this preliminary experiment, we identified low TBX3e-mCherry expression in two of the three clones (A4 and B5). Unfortunately, when repeating the BMP4 supplemented cardiac differentiation several times, we could not identify TBX3e-mCherry+ cells. As we initially found mCherry expression in the cardiac differentiation from clone A4, we decided to use this clone for further screening analysis, as it shows the best cardiac differentiation in presence of Xav939, compared to other clones (**Fig. 3b**).

Screening Protocols to Obtain Cardiac Conduction System Cells *In Vitro*

A number of signalling pathways has been shown, either *in vivo* or *in vitro*, to play a role in CCS development and/or differentiation[8, 13-17]. Here, we investigated the potential role of various factors in cardiac conduction lineage differentiation from hPSCs. First, we tested the potential role of neuregulin (NRG)-1β/ErbB signalling in cardiac conduction lineage differentiation, as it was previously shown that inhibition of this pathway resulted in enhanced differentiation towards a cardiac nodal phenotype[14]. However, addition of AG1478, an ErbB antagonist, from day 3-5, or day 5-7 during our standard cardiac differentiation protocol, did not result in activation of TBX3e-mCherry positive cells at any time point studied (**Fig. 3c**). Furthermore, the presence of non-canonical Wnt-agonist WNT11 was tested during differentiation, shortly after mesoderm formation (**Fig. 3d**). WNT11 is detected in developing central conduction structures in mouse, including the right atrioventricular ring[17]. Nonetheless, we did not identify TBX3e-mCherry positive cells. However, in a preliminary experiment we did observe that WNT11 could possibly enhance cardiac differentiation, indicated by increased NKX2-5-GFP+ levels in both conditions (D3-D5 and D5-D7), This is in agreement with previous studies[18-21]. Further, we tested a protocol that was recently developed by Birket et al.[22], which resulted in differentiation to pacemaker-like CMs in a transgenic hESC line. Growth factors involved in the derivation of NKX2-5 negative beating CMs from hESCs included SU5402 (an FGF-receptor-inhibitor[22], and BMP4. Additional protocols were screened, which also included AG1476. Birket et al. showed how inhibition of FGF signalling could potentially play a role in SAN lineage specification[22]. However, in our experiments, addition of both FGF-inhibitor SU5402 and BMP4 did not result in TBX3e-mCherry positive cells, and even inhibited cardiac differentiation, although not significantly, in the absence of XAV939 (**Fig. 3e**). Birket et al. also showed this inhibitory effect of SU5402 on NKX2-5 expression. In addition, the differentiations that also included AG1476 showed total absence of NKX2-5 positive CMs (**Fig. 3e**), indicating a role for neuregulin (NRG)-1β/ErbB signalling in cardiac differentiation, which is reviewed elsewhere[23].

DISCUSSION

In vivo developmental studies from different species are contributing largely to our increased understanding of combinatorial actions of regulatory elements on the precise patterned expression of specific cardiac genes, such as *Tbx3*[8]. Moreover, advancements in new technologies to understand chromatin organization and epigenetic changes upon lineage commitment, contribute to the identification of these critical regulatory elements[24, 25]. This increased knowledge on molecular regulation during the different phases of cardiac development can be extrapolated and applied in human *in vitro* differentiation models, by the use of fluorescent reporters in combination with medium- up to high-throughput screening methods. This will allow culture, identification and characterization of cardiac subtype cells of interest. The derivation of these subsets of cells will allow a next step in drug development: cardiac subtype-specific and disease-specific drug screenings.

The development of protocols to differentiate towards cells from the CCS will create opportunities for their use in the development of safe and effective drugs for the treatment of diseases specifically affecting the conductive system, including atrioventricular nodal abnormalities and may eventually lead to the production of biopacemakers. In this study, we describe the generation of a TBX3-reporter in the previously published NKX2-5-eGFP hESC line[9]. TBX3 is a cardiac transcription factor, crucial for the formation of the cardiac conduction system in the developing mammalian heart. Recently, the group of Christoffels identified, using high-resolution 4C-sequencing and transgenic mice models, 2 highly conserved enhancer regions within the Tbx3 domain. Both enhancers synergize with the Tbx3 promoter to drive TBX3 expression in the atrioventricular conduction system[8]. Here, we coupled these enhancers, eA and eB, to minimal promoter Hsp68 and mCherry coding sequences, and transfected hESCs with lentivirus particles containing the generated reporter cassette. We hypothesised that activation of both enhancers during induction of cardiac differentiation would result in mCherry expression in cell types that are recapitulating the TBX3 activity in the AV node *in vivo*. Subsequently, mCherry-positive, GFP-positive, nodal cardiomyocytes could be isolated and further characterized.

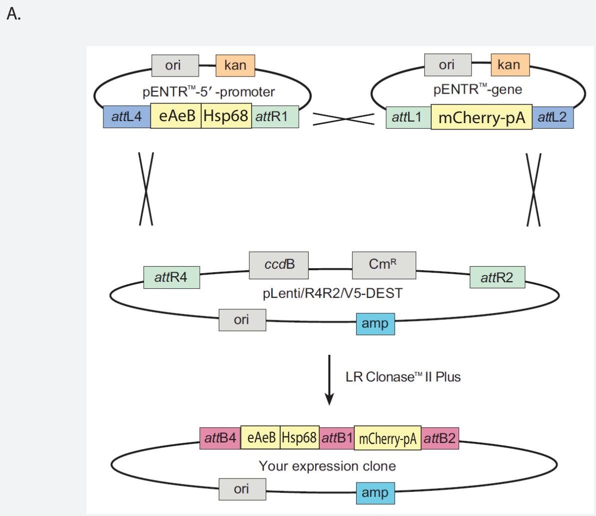
Extrapolating developmental processes from *in vivo* mouse studies to *in vitro* human differentiation models has been shown very useful due to the highly conserved regions that are shared between mouse and human. Moreover, comparison to the developmental processes and molecular mechanisms in other species is a necessity, since human material at early stages of development are extremely difficult, if not impossible, to obtain. However, our study appeared to be challenging due to several technical and biological difficulties, including the random integration properties of lentiviral particles, and limited information regarding the signalling pathways involved in cardiac conduction system lineage development. Upon lineage commitment, chromatin dynamics and epigenetic regulators may change, which may result in the risk of transgene silencing of the random integrated reporter cassette[26-28]. To circumvent this problem in first case, we decided to screen several clones on their potential

TBX3e-mCherry expression upon standard cardiac differentiation conditions, and in presence of BMP4. However, after an initial positive result, displaying a modest upregulation of mCherry expression in two clones, none of the clones showed mCherry expression in follow-up experiments. Moreover, this also indicated the lack of active eA and eB in our standard cardiac differentiation protocol. This is in line with previous findings on the low TBX3 expression levels upon cardiac differentiation[12]. We performed further screening differentiations on one clone, that showed the initial positive result (clone A4).

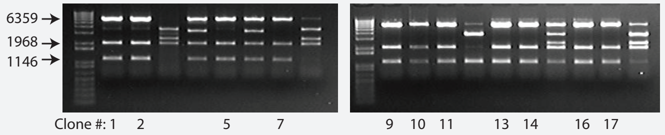
As van Weerd et al. showed eA and eB activation by the combinatorial presence of GATA4 and BMP4, we initially screened on TBX3e-mCherry expression by testing different concentrations of BMP4 at different timings of cardiac differentiation. Here, we assumed high endogenous expression levels of Gata4, based on previous gene expression array results[12]. Our nodal-like CMs are expected to co-express NKX2-5-eGFP and TBX3e-mCherry levels, as Nkx2-5 is required for the development and maintenance of the AVN[29]. Unfortunately, after the initial positive experiment, no TBX3e-mCherry positive cells could be identified in any condition. Based on knowledge from *in vivo* mouse development and recent publications on the *in vitro* derivation of human cardiac conduction cells[14, 22], growth factors were screened for their role in TBX3 expressing conduction cell derivation; a variety of differentiation protocols were developed. For instance, Birket et al. showed how FGF-signalling inhibition by SU5402 could have an inhibitory effect on NKX2-5 expression in early cardiac progenitors[22], which further differentiated to CMs with a pacemaker identity, based on electrophysiology and gene expression, including increased HCN4 and TBX18 levels. However, in our experiments, addition of SU5402 and BMP4 did not result in TBX3e-mCherry positive cells, and did even inhibit cardiac differentiation in the absence of XAV939. Further, Wnt11 and (NRG)-1β/ErbB signalling were both described to play a role in the developing cardiac conduction system[14, 17]. Again, also here, we could not detect any TBX3e-mCherry expression. Despite the absent of mCherry expressing cells, we could confirm the role of Wnt11 and neuregulin (NRG)-1β/ErbB signalling in heart development, as we did find increasing NKX2-5-eGFP expressing CMs when differentiation was induced with Wnt11, and decreasing levels after induction with ErbB antagonist AG1478. However, important to note is the high variance of biological repeats, which makes it more difficult to obtain statistically significant results. In order to make any conclusive statements on the effects of cytokine on cardiac differentiation and specification it is crucial to optimize and standardize cell culture and differentiations conditions in order to minimize biological variation.

Our results regarding the lack of induction of TBX3 activity is in contrast with previous findings from the studies of Zhu et al. and Birket et al. This could be due to 1) transgene silencing, or 2) restricted activity of the highly conserved eA and eB in only a subset of TBX3 expressing cardiac conduction cells (primordial AVN, AVC, AV rings, prospective AV bundle). The latter explanation is feasible, since both previously mentioned studies showed upregulation of TBX3 expression in their PSC-derived nodal cells[14, 22]. Indeed, TBX3 expression is critical for conduction cell specificity and function[6, 30, 31]. To circumvent the risk of transgene silencing due to random integration of the reporter cassette, the generation of an endogenous TBX3 fluorescent reporter line, using a knockin through replacement or fusion strategy could be recommended. However, as TBX3 expression is not specifically expressed in the heart, the use of specific regulatory elements, coupled to fluorescent reporters, or in combination with another cardiomyocyte reporter, such as a sarcomeric protein, or NKX2-5, such as in this study, would still be a preference. Targeted integration of a reporter cassette, such as used in our study, at an ubiquitous active genomic region ("a docking station") should be reconsidered and could make it more promising to generate such fluorescent reporter lines. Genomic sites that have been described in literature include including sequences highly homologous to the mouse Rosa26 locus and the AAVS1 locus[32-34].

FIGURES



B. Digestion of modfiied destination clones with EcoRV



C. Digestion of modified destination clones with PvuI

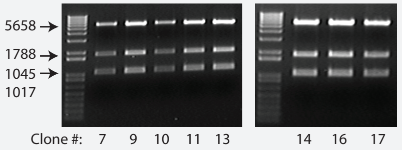
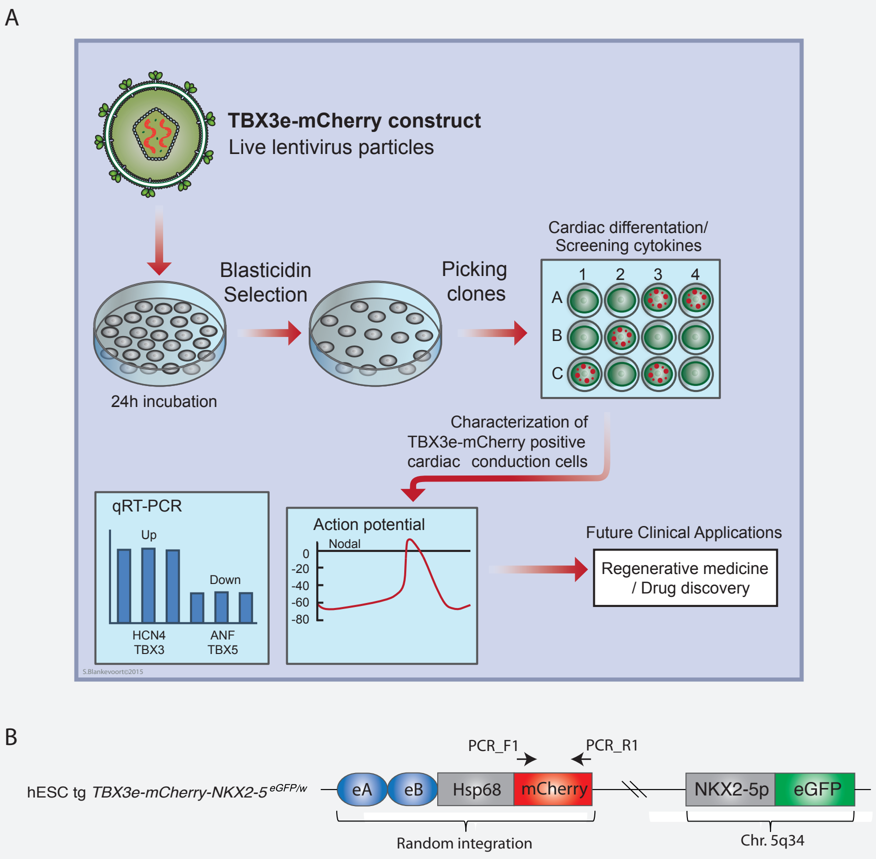


Figure 1. A: Gateway® cloning of the two TBX3 enhancers eA and eB and the Hsp68 promoter element, and mCherry from two separate entry vectors into one destination vector (Life Technologies). **B.** Digestion with EcoRV of multiple picked destination vectors to define correctly cloned vectors**. C.** Digestion with PvuI of a selection of correctly cloned destination vectors to confirm its correct integrity.



C. PCR screening after transfection: PCR_F1 + PCR_R1

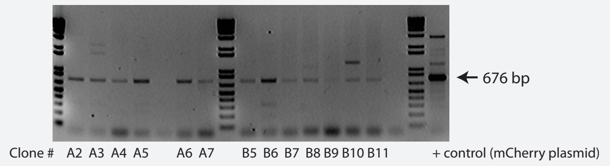
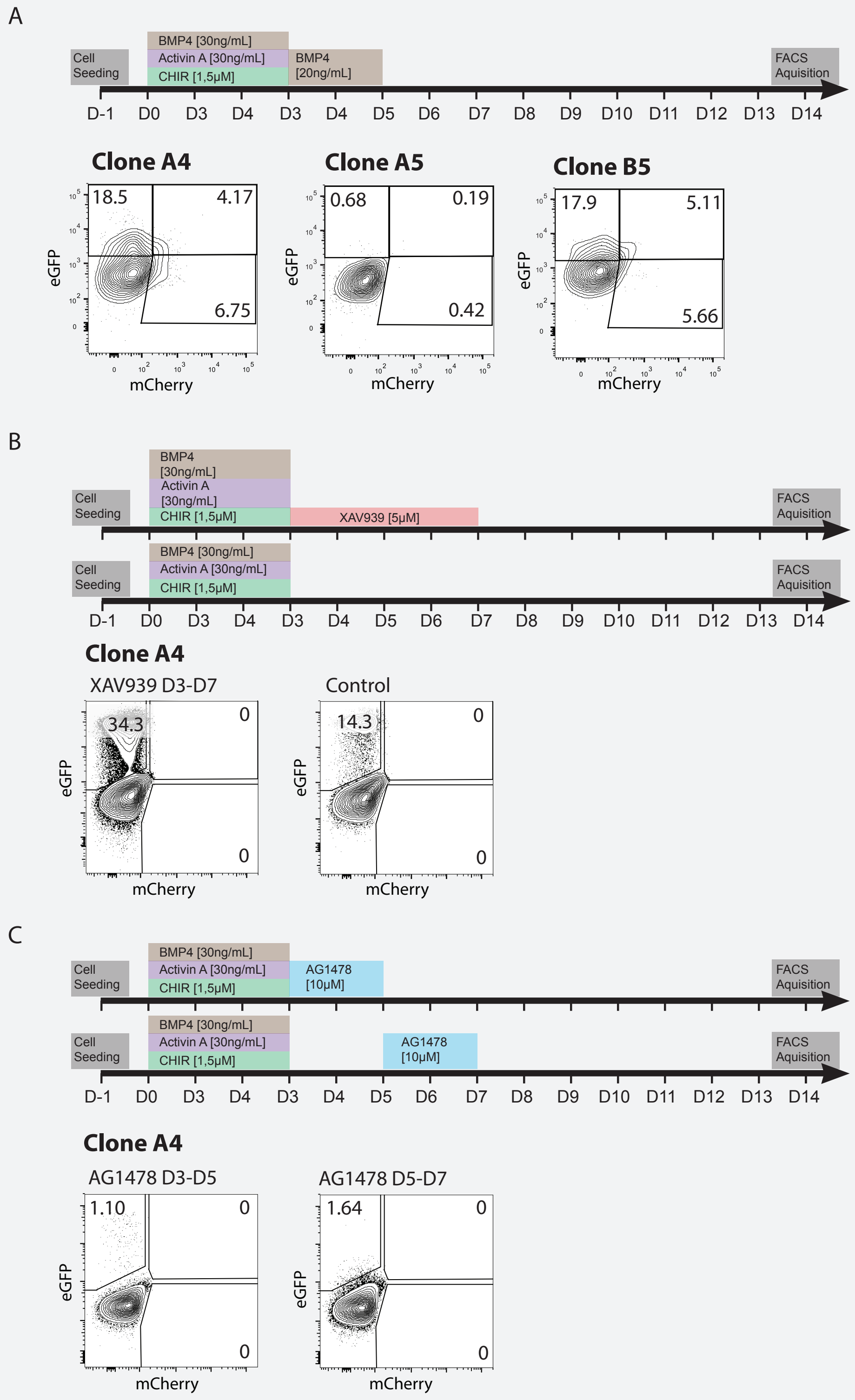


Figure 2. A: Generation of the TBX3e-mCherry hESC reporter line through viral delivery of the TBX3e-mCherry reporter construct. Cells were incubated with lentiviral particles for 24 hours followed by selection using blasticidin. Resistant clones were picked and screened for mCherry expression upon cardiac differentiation. TBX3e-mCherry expressing cardiomyocytes would be characterized on their gene expression profile and action potential phenotype, and could be of potential for future translational applications. **B:** Schematic overview of the genome of the dual cardiac reporter line. **C**. PCR screening of the targeted clones shows that several clones contain the integrated reporter cassette.



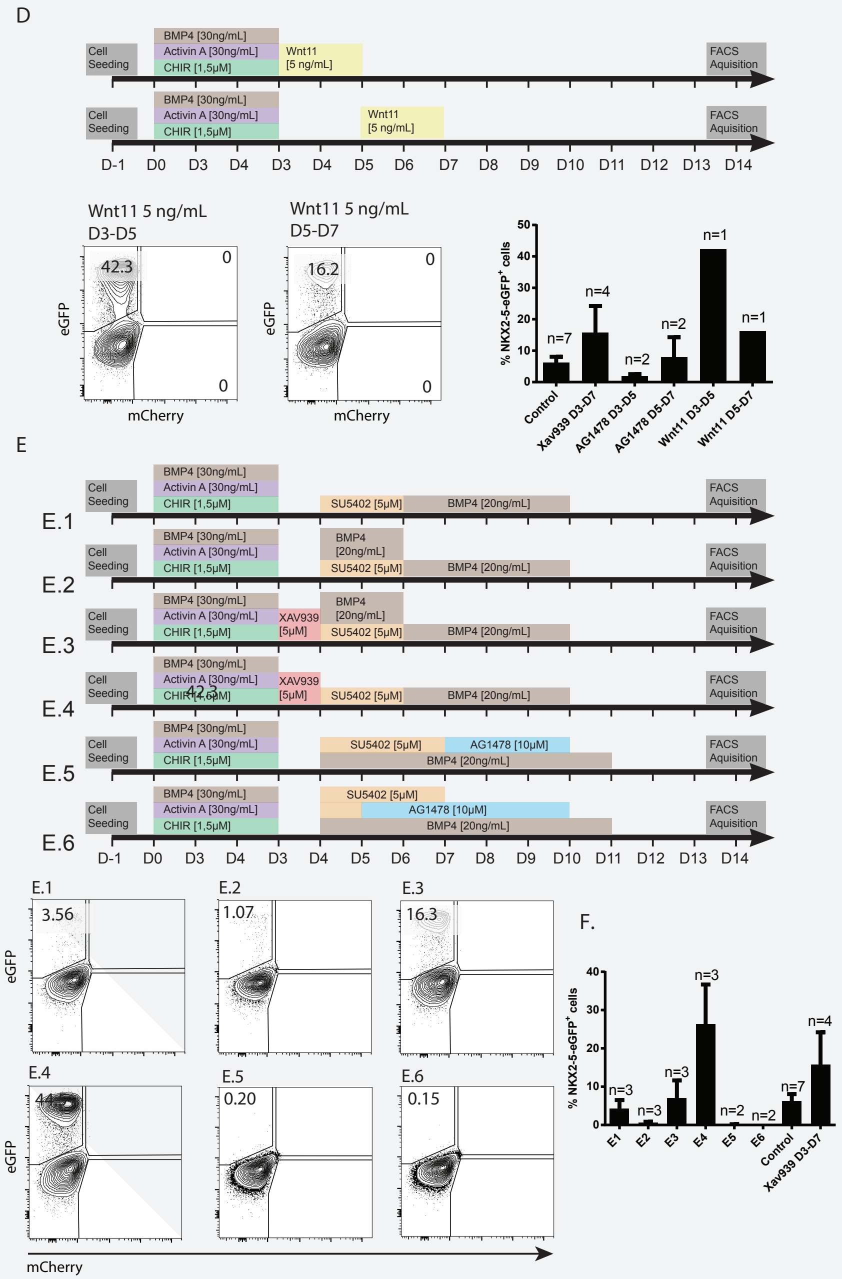
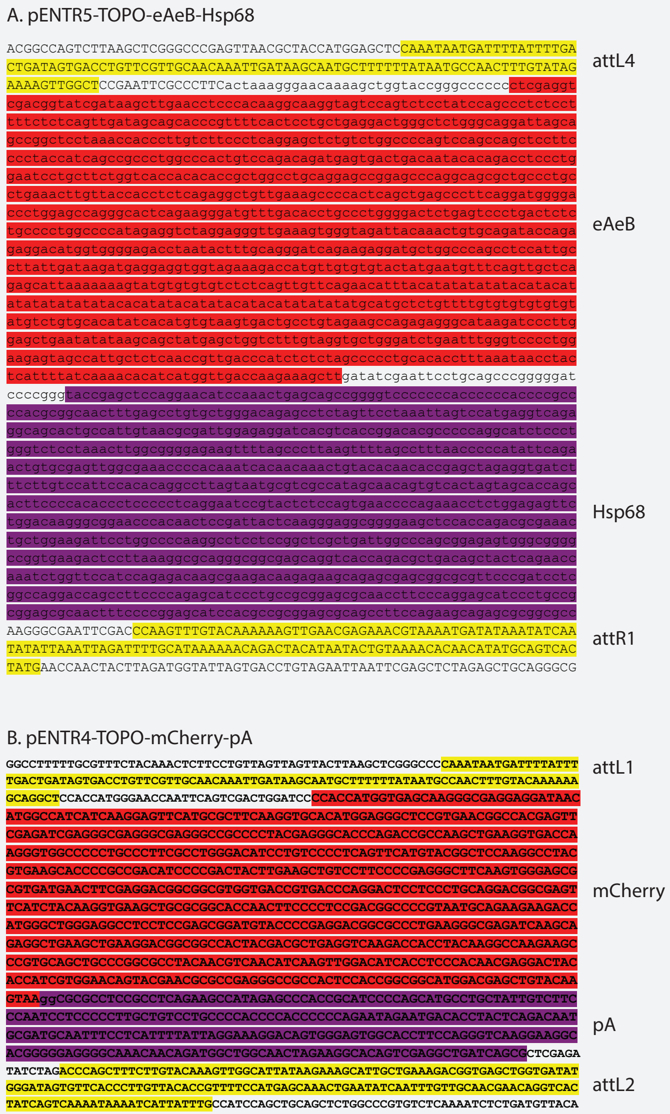


Figure 3. Growth factor screenings during cardiac differentiations in order to direct differentiation towards TBX3e-mCherry expressing cardiomyocytes. **A:** Addition of 20ng/mL BMP4 resulted in a small percentage of mCherry expressing cardiomyocytes at day 14 of differentiation in clone A4 and clone B5. **B:** A standard monolayer differentiation protocol results in 34% of NKX2-5-eGFP expressing cardiomyocytes at day 14, in clone A4, versus 14% without Wnt signalling inhibitor Xav939. **C:** Addition of ErbB antagonist AG1478 after mesoderm formation between D3-D5 or D5-D7 did not result in TBX3e-mCherry expressing cardiomyocytes at day 14 , and did even inhibit cardiac differentiation. **D:** Addition of Wnt11 after mesoderm formation did not result in TBX3e-mCherry expressing cardiomyocytes. Addition of Wnt11 between D3-D5 could possible enhance cardiac differentiation, but requires repeated experiments. **E,F:** Screening of a variant of differentiation protocols did not direct towards TBX3e-mCherry expressing cardiomyocytes.



Supplemental Figure 1. Sequences of pEntry vectors containing both TBX3 enhancers eA and eB and Hsp68, and mCherry, surrounded by aTT sites (Lambda attachment sites) for recombination.

REFERENCES

1. Christoffels VM, Smits GJ, Kispert A, et al. Development of the Pacemaker Tissues of the Heart. CIRCULATION RESEARCH 2010;106(2):240–254.
2. Bruneau BG, Nemer G, Schmitt JP, et al. A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. CELL 2001.
3. Olson EN. Gene Regulatory Networks in the Evolution and Development of the Heart. SCIENCE 2006;313(5795):1922–1927.
4. Hatcher CJ, Basson CT. Specification of the Cardiac Conduction System by Transcription Factors. CIRCULATION RESEARCH 2009;105(7):620–630.
5. HOOGAARS W. The transcriptional repressor Tbx3 delineates the developing central conduction system of the heart. CARDIOVASCULAR RESEARCH 2004;62(3):489–499.
6. Hoogaars WMH, Engel A, Brons JF, et al. Tbx3 controls the sinoatrial node gene program and imposes pacemaker function on the atria. GENETICS 2007;21(9):1098–1112.
7. Horsthuis T, Buermans HPJ, Brons JF, et al. Gene Expression Profiling of the Forming Atrioventricular Node Using a Novel Tbx3-Based Node-Specific Transgenic Reporter. CIRCULATION RESEARCH 2009;105(1):61–69.
8. van Weerd JH, Badi I, van den Boogaard M, et al. A large permissive regulatory domain exclusively controls tbx3 expression in the cardiac conduction system. CIRCULATION RESEARCH 2014;(115):432–441.
9. Elliott DA, Braam SR, Koutsis K, et al. NKX2-5eGFP/w hESCs for isolation of human cardiac progenitors and cardiomyocytes. NAT METH 2011;8(12):1037–1040.
10. Ng ES, Davis R, Stanley EG, et al. A protocol describing the use of a recombinant protein-based, animal product-free medium (APEL) for human embryonic stem cell differentiation as spin embryoid bodies. NATURE PROTOCOLS 2008;3(5):768–776.
11. Wiese C, Grieskamp T, Airik R, et al. Formation of the Sinus Node Head and Differentiation of Sinus Node Myocardium Are Independently Regulated by Tbx18 and Tbx3. CIRCULATION RESEARCH 2009;104(3):388–397.
12. Hartogh Den SC, Schreurs C, Monshouwer-Kloots JJ, et al. Dual Reporter MESP1 mCherry/w-NKX2-5 eGFP/whESCs Enable Studying Early Human Cardiac Differentiation. STEM CELLS 2014;33(1):56–67.
13. Sun C, Yu D, Ye W, et al. The short stature homeobox 2 (Shox2)-bone morphogenetic protein (BMP) pathway regulates dorsal mesenchymal protrusion development and its temporary function …. JOURNAL OF BIOLOGICAL … 2015.
14. Zhu W-Z, Xie Y, Moyes KW, et al. Neuregulin/ErbB Signaling Regulates Cardiac Subtype Specification in Differentiating Human Embryonic Stem CellsNovelty and Significance. CIRCULATION … 2010;(107):776–786.
15. Notch signaling reprograms cardiomyocytes to a conduction-like phenotype. CIRCULATION … 2012;(126):1058–1066.
16. Milan DJ. Notch1b and neuregulin are required for specification of central cardiac conduction tissue. DEVELOPMENT 2006;133(6):1125–1132.
17. Bond J, Sedmera D, Jourdan J, et al. Wnt11 and Wnt7a are up-regulated in association with differentiation of cardiac conduction cells in vitro and in vivo. DEV. DYN. 2003;227(4):536–543.
18. Pandur P, Läsche M, Eisenberg LM, et al. Wnt-11 activation of a non-canonical Wnt signalling pathway is required for cardiogenesis. NATURE 2002.
19. Bisson JA, Mills B, Paul Helt J-C, et al. Wnt5a and Wnt11 inhibit the canonical Wnt pathway and promote cardiac progenitor development via the Caspase-dependent degradation of AKT. DEVELOPMENTAL BIOLOGY 2015;398(1):80–96.
20. Cohen ED, Miller MF, Wang Z, et al. Wnt5a and Wnt11 are essential for second heart field progenitor development. DEVELOPMENT 2012;139(11):1931–1940.
21. Nagy II, Railo A, Rapila R, et al. Wnt-11 signalling controls ventricular myocardium development by patterning N-cadherin and beta-catenin expression. CARDIOVASCULAR RESEARCH 2010;85(1):100–109.
22. 22. Expansion and patterning of cardiovascular progenitors derived from human pluripotent stem cells. NATURE BIOTECHNOLOGY 2015;33(9):970–979.
23. Odiete O, Hill MF, Sawyer DB. Neuregulin in cardiovascular development and disease. CIRCULATION RESEARCH 2012;111(10):1376–1385.
24. Wamstad JA, Alexander JM, Truty RM, et al. Dynamic and Coordinated Epigenetic Regulation of Developmental Transitions in the Cardiac Lineage. CELL 2012;151(1):206–220.
25. Paige SL, Thomas S, Stoick-Cooper CL, et al. A Temporal Chromatin Signature in Human Embryonic Stem Cells Identifies Regulators of Cardiac Development. CELL 2012;151(1):221–232.
26. J BURKE les, BANIAHMAD A. Co-repressors 2000. THE FASEB JOURNAL 2000;14(13):1876–1888.
27. . MacArthur CC, Xue H, Van Hoof D, et al. Chromatin Insulator Elements Block Transgene Silencing in Engineered Human Embryonic Stem Cell Lines at a Defined Chromosome 13 Locus. STEM CELLS AND DEVELOPMENT 2012;21(2):191–205.
28. Xia X, Zhang Y, Zieth CR, et al. Transgenes Delivered by Lentiviral Vector are Suppressed in Human Embryonic Stem Cells in A Promoter-Dependent Manner. STEM CELLS AND DEVELOPMENT 2007;16(1):167–176.
29. Jay PY, Harris BS, Maguire CT, et al. Nkx2-5 mutation causes anatomic hypoplasia of the cardiac conduction system. J. CLIN. INVEST. 2004;113(8):1130–1137.
30. Bakker ML, Boink GJJ, Boukens BJ, et al. T-box transcription factor TBX3 reprogrammes mature cardiac myocytes into pacemaker-like cells. CARDIOVASCULAR RESEARCH 2012;94(3):439–449.
31. Frank DU, Carter KL, Thomas KR, et al. Lethal arrhythmias in Tbx3-deficient mice reveal extreme dosage sensitivity of cardiac conduction system function and homeostasis. PNAS 2012;109(3):E154–E163.
32. Costa M, Dottori M, Ng E, et al. The hESC line Envy expresses high levels of GFP in all differentiated progeny. NAT METH 2005;2(4):259–260.
33. Irion S, Luche H, Gadue P, et al. Identification and targeting of the ROSA26 locus in human embryonic stem cells. NATURE BIOTECHNOLOGY 2007;25(12):1477–1482.
34. . Hockemeyer D, Soldner F, Beard C, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. NATURE BIOTECHNOLOGY 2009;27(9):851–857.

CHAPTER 6

A Novel Approach for the Pure Isolation of Ventricular Cardiomyocytes from human Embryonic Stem Cells: a dual MYL2-T2A-mCherry NKX2-5-eGFP Fluorescent Reporter Line

Sabine den Hartogh[1], Chantal Schreurs[1], Catarina Grandela[2],
Christine Mummery[1], Robert Passier[1,3]

1. Leiden University Medical Center, Dept. Anatomy and Embryology, Leiden, The Netherlands
2. Pluriomics BV, Bioscience Park, Leiden, The Netherlands.
3. Department of Applied Stem cell Technologies. MIRA Institute for Biomedical Technology and Technical Medicine. University of Twente, P.O.Box 217, Enschede, The Netherlands.

ABSTRACT

In recent years, differentiation of human pluripotent stem cells (PSCs) to the cardiac lineage has improved considerably leading to mixtures of cardiomyocyte subtypes. This allowed researchers to study human cardiac development and differentiation, and to obtain cardiomyocytes (CMs) for pre-clinical applications, such as drug discovery, safety pharmacology, disease modelling and cell-based therapies for cardiac repair. However, in order to increase the predictability and the robustness of stem cell based assays and follow-up experiments, it is essential to obtain pure populations of CM subtypes. Here, we describe the generation of a MYL2-T2A-mCherry-NKX2-5-eGFP hESC reporter line in order to study and steer *in vitro* cardiac differentiation to a ventricular cardiomyocyte subtype. MYL2 encodes for ventricular myosin light chain-2 (MLC-2v) and is specifically expressed in ventricular cardiomyocytes. During cardiac differentiation, MYL2 only becomes markedly expressed after 1 month of differentiation. Fluorescent labelling of cells expressing MYL2 will allow us to isolate ventricular cardiomyocytes and study their molecular and electrophysiological characteristics. Moreover, this may contribute to the identification of regulators that are important for the formation of the ventricular-specific-lineage. In this chapter, we focus on the technical challenges that we faced upon the generation of this fluorescent reporter line and our preliminary results on the identification of MYL2-T2A-mCherry expressing cardiomyocytes during cardiac differentiation.

INTRODUCTION

The heart is one of the first organs formed in the developing human embryo, indicating its vital importance in supplying oxygen and nutrients to the developing organism and at the later adult stage[1]. Formation of the developing heart is complex and requires a balanced interplay between specific molecular signalling pathways and transcription factors under tight temporal and spatial control. Congenital heart disease encompasses abnormalities in the developing heart due to mutations in important regulators of heart development, but may also be caused by prenatal environmental influences[2]. It is important to understand how these underlying defects result in the onset of congenital heart disease, which requires a complete understanding of the developing human heart. Acquired heart diseases, such as myocardial infarction, mostly result in a massive loss of ventricular cardiomyocytes (CMs) due to an ischemic event, leading to decreased cardiac function[1]. As cardiac disease is still a leading cause of morbidity and mortality worldwide, there is an urgent need for a better understanding of its onset and progression, and in the end, better treatment options[3, 4]. The adult human heart has no or limited regenerative capacity, and is therefore unable to restore damaged cardiac cells due to ischemia or cardiomyopathies, resulting in progressive heart failure[5]. Replacement of damaged heart tissue with viable hPSC-derived cardiomyocytes may be a possible therapeutic strategy. In particular, the possibility to generate human iPSC-derived cardiomyocytes created the opportunity to use cardiomyocytes for autologous transplantation, as opposed to the ethically sensitive use of human embryonic stem cells (hESCs). Optimized hPSC-CM protocols have been developed and result in 50-95% CM differentiation efficiencies[6]. Moreover, the use of hPSC-CMs offers an exclusive physiologically relevant human model system for drug discovery and pre-clinical development strategies. However, current cardiac differentiation protocols result in heterogeneous pools of CMs. Although these mixed cultures predominantly consist of ventricular-like cells, there is a significant percentage of other cardiomyocyte subtypes, such as atrial-like and nodal-like cells[7]. For high reliability in predictive pharmacological compound screenings, pure populations of cardiac subtypes are necessary. In particular, ventricular cardiomyocytes are of interest, since in the majority of cases, genetic cardiac disorders and cardiac injury affect ventricular cardiomyocytes.

The generation of fluorescent pluripotent stem cell reporter lines offer an excellent opportunity to study the derivation of specific cellular subtypes[8, 9]. In order to obtain pure populations of ventricular CMs, here, we generated a MYL2 (MLC2V) fluorescent reporter in the previously described NKX2-5-eGFP hESC line[9]. *In vivo*, MLC2V expression is restricted to the ventricular segment of the developing human heart, followed by expressed in the entire ventricular compartment in the adult heart[10]. Previous studies have described the generation of MYL2 reporter lines in mouse ESCs, reporting a specific MYL2 promoter element or endogenous MYL2 expression[11, 12]. Here, we reported endogenous MYL2 expression in hESCs through a fusion strategy, making use of a cleaving 2A peptide sequence which is positioned in frame between MYL2 and the coding sequence of mCherry[13]. 2A-mediated cleavage is a universal phenomenon in all eukaryotic cells. However, it has not been observed in prokaryotic cells[14]. Transcriptional activity of the MYL2 transcript leads to production of a single messenger RNA encoding both proteins, MYL2 and mCherry. Here, we will discuss our technical approach for the generation of a MYL2-fluorescent reporter line, preliminary results, and possible future strategies for obtaining pure ventricular CM populations.

MATERIAL AND METHODS

Generation of the MYL2 targeting construct

To generate the *MYL2* targeting vector, bacterial artificial chromosome RP11-587H5, containing the MYL2 sequence (Source Bioscience, Lifesciences, Nottingham, UK), was modified by recombineering (Gene Bridges, Heidelberg, Germany) **(Fig. 1)**. An R6k plasmid was generated through cloning, containing the T2A sequence[15], followed by mCherry (in-frame), and a neomycin selection cassette, flanked by LoxP recombination sites. 20ug of the complete cassette of interest was linearized by the two enzymes PvuI and EcorV, restricting an upstream PvuI site, and a downstream EcoRV site **(Fig. 1a)**. Recombineering was performed using the Red/ET recombination technology (Genebridges). Here, we introduced the Red/ET plasmid (pSC101-BAD-gbaA-tet) into E. Coli containing BAC RP11-587H5 through electroporation. Genes from plasmid pRed/ET promote base precise exchange of DNA sequences flanked by homology arms. The exonuclease Redα and the DNA annealing protein Redβ catalyze the *in vivo* reaction. Further, two primers were designed, containing 40bp that were homologous to the ends of the linearized R6K cassette, and containing 50bp homologous to the region of the MYL2 BAC upstream and downstream of the MYL2 stopcodon (**Fig 1b**). Addition of 10% L-arabinose (to a final concentration of 0.1%) for induction of RedET genes, followed by electroporation of the linearized R6k cassette and both primers into the MYL2 BAC containing E. Coli, resulted in quadruple recombination of the BAC. Oligonucleotide primers oligo1 and oligo2 were used for recombineering (**Table 1**). Recombineered BACs were screened by PCR; 7 out of 40 clones displayed a correct product size (17.5%)(**Fig. 1c**). For clone 3, PCR products were generated for the 3' homology arm (HA) and 5' HA recombined regions, and sent for sequencing (100% correct). The modified MYL2 locus, including 5' and 3' homology arms, were subsequently subcloned into a MYL2 targeting vector. For this, a PCR (Phusion, Thermo Scientific) was performed on a linearized minimal vector (digested by XhoI and HindIII), containing a Diphtheria Toxin A (DTA) sequence for negative selection, an ampicillin resistance cassette, and an origin of replication (ORI) sequence (**Fig. 2a**). Oligos that are used to generate this PCR product contain extra 50bp sequences that are homologous to the BAC sequence that are upstream (5' HA, 8.8kb) and downstream (3' HA, 2.3kb), so that regions in between these homologous regions are subcloned into a final target construct, containing the PCR product as backbone (**Fig. 2a, Table 1**). For subcloning, the pRed/ET plasmid was re-introduced into the E.Coli, containing modified BAC clone 3, and activated by L-arabinose prior to electroporation of the PCR product of the minimal vector. The newly subcloned minimal vectors were then isolated and screened by PCR; 80% of the clones was positive (**Fig. 2b**). Three clones were further verified using Bgl1 and XbaI digestions (**Fig. 2c**). Clone 3.9 displayed the correct sequence and was further used for MYL2 targeting.

Generation of the MYL2-T2A-mCherry Reporter Line

The final targeting vector comprised a 8.8-kb 5_homology arm, T2A-mCherry, loxP-flanked G418 resistance cassette, a 2.3-kb 3_homology arm, and a Diphtheria Toxin A (DTA) sequence for negative selection (**Fig. 2a**). The targeting vector was linearized by PvuI digestion and 20 µg was electroporated into 10x10^6 NKX2-5-eGFP reporter hESCs (HES3-NKX2-5eGFP/w) [9] as described previously[16]. Homologous recombination will result in replacement of the MYL2 STOP codon, localized in exon 7. Targeted clones were identified by PCR using oligonucleotides specific for mCherry (MYL2_screen_F1: CCCCCTGAACCTGAAACATA) and a region immediately 3' of the targeting vector (MYL2_screen_R1: GCATCTCTGTATGTCCCTGTG) (**Fig. 3a,b**). After identification of correctly targeted clones, five hESC clones were maintained as single cell cultures, enzymatically passaged, and transfected with a Cre recombinase construct using lipofectamin (Invitrogen) for removal of the neomycin cassette. 24 Hours after transfection, cells were cultured in the presence of puromycin (1ug/mL) for 36 hours. After puromycin selection, single cells were sorted based on SSEA-4 expression, screened for neomycin excision, and further cultured as single cells (**Fig. 3a,c**).

By the use of a T2A cleavage peptide to report MYL2 gene activity, endogenous gene expression remains normal. Here, we replaced the MYL2 stopcodon with the T2A and mCherry sequences (in-frame), and a selection cassette, which is excised after targeting to prevent interference with transcriptional activity. Upon MYL2 activation, one RNA transcript is generated, followed by ribosomal translation, where T2A is cleaved through ribosomal skipping, resulting in the presence of two separate proteins MYL2 and mCherry (**Fig. 3e**).

Flow Cytometry

For analysis on MYL2-T2A-mCherry and NKX2-5-eGFP expression, cells were dissociated using 1x or 10x TrypLe Select (Life Technologies). After dissociation, cells were resuspended in FACS Wash buffer (PBS, 2% FCS (Gibco), 2mM EDTA) and filtered through a 40µm cell strainer (Falcon). Cells were analyzed using a Miltenyi VYB Flow Cytometer using a 488 nm and a 561 nm laser.

Single Cell Sorting:

For single cell sorting, hESCs were enzymatically passaged using 1x TrypLE Select (Life Technologies). After dissociation, cells were resuspended in FACS sorting buffer (PBS with 0.5% BSA and 2.5mM EDTA) and filtered through a 40µm cell strainer (Falcon). Cells were sorted using a BD ARIA III flow cytometer into 96 wells flat bottom low attachment plates. Live cells were gated on the basis of side scatter, forward scatter and propidium iodide exclusion. Flow cytometric gates were set using control cells labelled with the appropriate isotype control antibody.

Cardiac differentiation

MYL2-mCherry/w-NKX2-5-eGFP/w hESCs were cultured on mouse embryonic fibroblasts in 2.85 cm2 organ dishes (Falcon), in 1 mL hESC medium (DMEM F12, Non Essential Amino Acids, Knock Out Serum (Gibco by Life Technologies); bFGF (Miltenyi Biotec)) and were mechanically passaged. One or two days before differentiation, cells were passaged and seeded in a 12-well tissue culture dish coated with Matrigel (BD Biosciences; growth factor reduced, phenol red–free), in 1 mL of hESC medium per well. On the first day of differentiation, day 0, hESC medium was replaced by low insulin serum-free medium BPEL (BSA, polyvinyl alcohol, essential lipids, as previously described[9] containing the cytokines BMP4 (30 ng ml−1, R&D Systems) and Activin A (30 ngml−1, Miltenyi Biotec), and Chir99021 (1.5uM, AxonMedchem). At day 3 of differentiation, the medium was refreshed with BPEL containing 5uM Xav939 (R&D). On day 7 of differentiation, the medium was replaced with BPEL without growth factors.

To enhance MYL2 expression, cardiomyocytes were dissociated at day 21 of differentiation with 10x TrypLe (Gibco) for 10 minutes at room temperature, and plated as single cells in Cardiomyocyte Medium (CM) (Pluriomics BV). After 7 days, medium was changed to commercially available maturation medium (MM), containing T3 hormone (Pluricyte medium, Pluriomics BV).

RESULTS

Generation of the Dual Cardiac Reporter hESC Line MYL2-T2A-mCherry-NKX2-5-eGFP

In order to isolate and characterize ventricular-specific cardiomyocytes, derived from hPSCs, we generated a MYL2 fluorescent reporter line in the previously generated NKX2-5-eGFP hESC line[9]. The MYL2 targeting construct was generated, linearized and electroporated into NKX2-5-eGFP hESC line (**Fig3a**). Homologous recombination resulted in the replacement of the MYL2 stopcodon with the T2A-mCherry sequences and a selection cassette (**Fig3a**). Correctly targeted clones were identified by PCR screening (**Fig 3a, b, c**). Five correctly targeted clones were obtained (18% targeting efficiency), transfected with Cre recombinase for excision of the neomycin-resistance cassette, and after selection of 36 hours, sorted as single cells. From these clones, one clone (8.1D1) that showed neomycin excision was selected for further experiments in this study (**Fig 3c**). A schematic overview of the newly generated dual reporter line is given in Figure 2a.

Ventricular Cardiomyocyte Derivation from hESCs

To study the presence of MYL2-mCherry expressing cells upon cardiac differentiation from hESCs, we analysed cardiac cells on mCherry and eGFP expression at different timepoints during differentiation (**Fig. 4a**). In order to increase expression of MYL2 in cardiomyocytes, cells were dissociated at day 21 of differentiation and further maintained in physiologically optimized CM medium (*Pluriomics, BV*). 7 Days after dissociation (day 28 since the start of differentiation), we analysed MYL2-T2A-mCherry expression by FACS. We observed that ~30% of cells expressed NKX2-5-eGFP levels at day 28 of differentiation. However, only about ~3% from the total differentiated cell population was expressing MYL2-T2A-mCherry, which was ±16% from total NKX2-5-eGFP expressing cardiac cells (**Fig. 4b**). It has been previously shown that MYL2 expression is only strongly increased after approximately one month of cardiac differentiation[12]. Therefore, we extended cardiomyocyte differentiation in the presence of maturation-promoting media (MM). At day 56 of differentiation, cells were analysed by FACS (**Fig. 4b**). When we cultured dissociated CMs in CM medium, or additionally in MM medium, we detected a decreased percentage of NKX2-5-eGFP expressing cells (~7-17%), which is most likely due to the proliferative capacity of non-cardiomyocytes that are present in these differentiation cultures. Nonetheless, we found a clear increase in the percentage of MYL2-T2A-mCherry positive cells in the NKX2-5-eGFP positive populations (~27-39%), indicating an increase in the proportion of ventricular cardiomyocytes in prolonged and optimized culture conditions (**Fig 4b**). In agreement, we could confirm the presence of double NKX2-5-eGFP+MYL2-T2A-mCherry+ cardiomyocytes by fluorescent microscopy, showing endogenous mCherry expression in day 67 dissociated cardiomyocytes (**Fig 4c**). Interestingly, when we measured NKX2-5-eGFP+MYL2-T2A-mCherry+ levels in undissociated cultures at day 56 of differentiation, which were continuously cultured in BPEL medium, we could almost not detect any MYL2-T2A-mCherry expressing cells (~0.2%) (**Fig. 4b**). Thus, either by dissociating cardiomyocytes, and thereby inducing mechanical stress, or by culturing in medium containing maturation-inducing-components, or both, may result in MYL2 upregulation, allowing detection and isolation of MYL2-T2A-mCherry positive cells.

DISCUSSION

Here, we describe the generation of a MYL2-fluorescent reporter line for the identification of hESC-derived ventricular CMs, in order to allow isolation and to further study ventricular-specific CMs. We developed a targeting strategy in which one allele of MYL2 was genetically modified into a bicistronic site, by replacing the stop codon of MYL2 for the coding sequences for a T2A viral peptide and fluorescent protein mCherry. When transcription of MYL2 is activated, one single messenger RNA molecule is transcribed. Upon ribosomal translation, T2A is cleaved, resulting in the presence of the two separate proteins MYL2 and mCherry. Small peptide T2A is described to be more reliable and efficient than the currently used internal ribosomal entry site (IRES) and leads to the expression of multiple cistrons at equimolar levels[13, 17-19]. Moreover, the use of a multi-cistronic fusion reporter construct, can be superior to a knockin replacement strategy in which one allele of the gene of interest is replaced by the fluorescent reporter construct, resulting in haplo-insufficiency and thus differential expression levels of the gene.

Here, we show that upon cardiac differentiation, up to day 28, only low levels of MYL2 expressing cardiomyocytes could be detected. This is in alignment with previous cardiac differentiation studies in mESCs, where also only 0.5 - 1.2 % of MYL2+ expressing cells could be identified between day 10 and day 30 of differentiation[11, 12].

This is in significant contrast with the presence of high numbers of ventricular CMs in current cardiomyocyte cultures, based on electrophysiology data[20], indicating the preference to search for an alternative ventricular marker for the efficient isolation of early ventricular CMs during differentiation. Only one earlier study in hESCs reports the presence of 67-98% MYL2-eGFP positive cardiomyocytes, but this has not been repeated[21]. Here, they made use of a randomly integrated reporter construct, which expression could have been influenced by surrounding genomic regulatory elements.

Further, we show that either dissociation of cardiac monolayers into single cell cultures or culturing cardiomyocytes in thyroid hormone-based culture medium additionally enhances MYL2 levels. Thyroid hormone is described to be a critical regulator of cardiac growth and development, both in fetal life and postnatally, and induces mature electrophysiological properties[12, 22-24]. In our study, it will be of further interest to isolate the double positive MYL2-T2A-mCherry-NKX2-5-eGFP cardiomyocytes, and study their electrophysiological and molecular properties, including genome wide analysis for the identification of ventricular-specific cell surface markers and important molecular regulators. The identification of a ventricular-specific cell surface marker, would allow ventricular CM isolation from any PSC-derived cardiac culture, without the need of genetic modifications. Here, we show that, although MYL2 could potentially be a proper marker for ventricular CM isolation, due to its low expression in early ventricular CMs, the identification of an additional early ventricular-specific marker would be a preference for the isolation of (early) ventricular CMs. Moreover, as MYL2 expression levels increase upon further maturation of our cardiomyocytes, it could be a potential maturation marker.

FIGURES

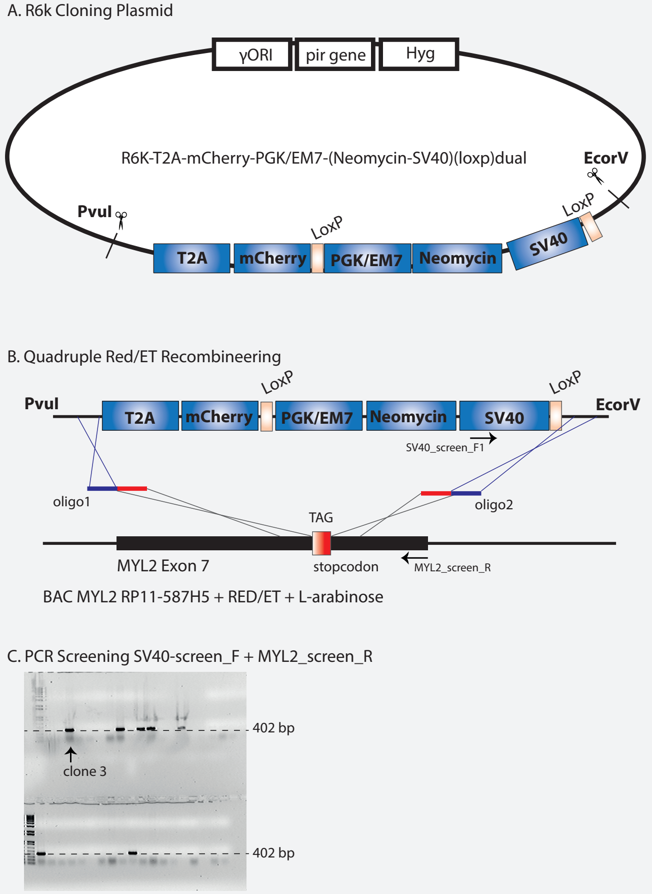
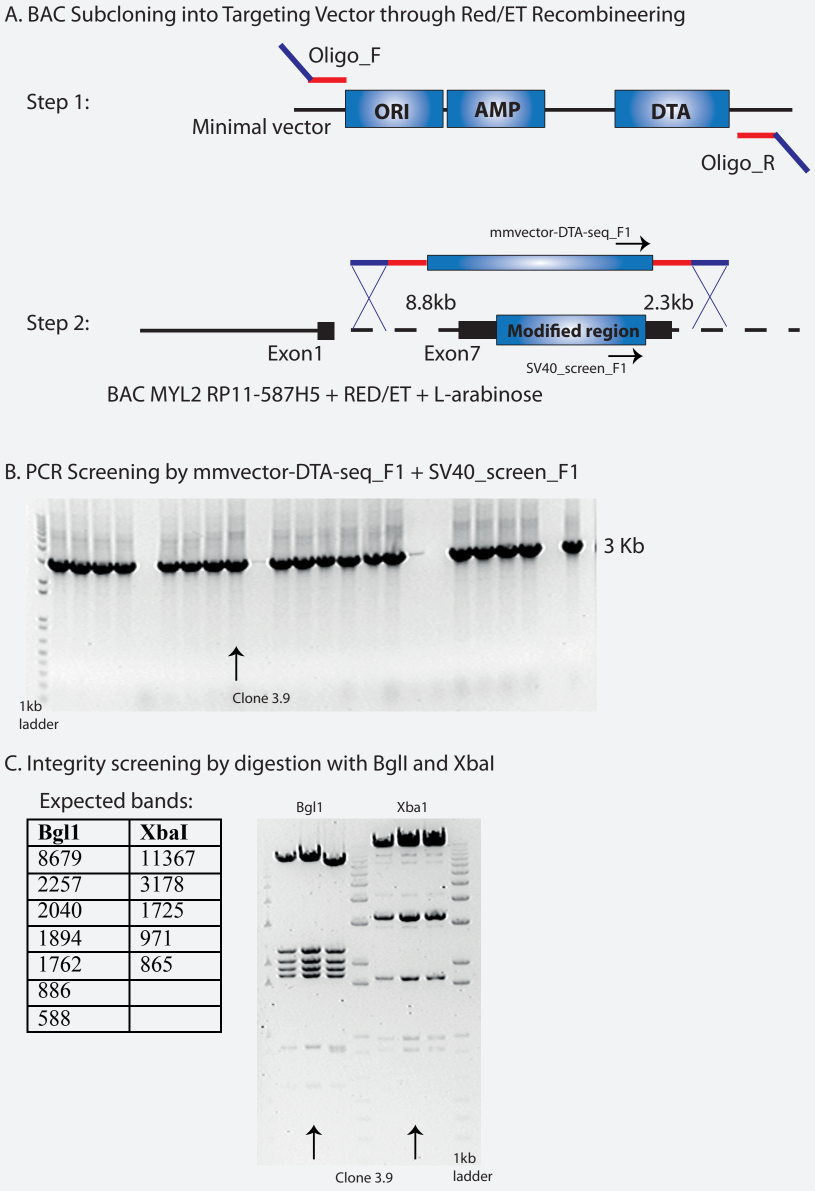


Figure 1. MYL2 BAC recombineering. **A:** An R6k cloning plasmid containing the T2A sequence, in-frame with the following mCherry sequence, and a eukaryote/bacterial selection cassette, was digested with PvuI and EcorV to obtain a linearized construct. **B:** A BAC, containing the MYL2 coding sequence, was modified through Red/ET quadruple recombineering (GeneBridges BV) T2A, mCherry, and the selection cassette replaced the TAG stopcodon sequence of MYL2, located in its last exon 7. **C:** PCR screening on BAC clones to identify for correctly modified BAC.



Bgl1	XbaI
8679	11367
2257	3178
2040	1725
1894	971
1762	865
886	
588	

Figure 2. BAC subcloning into targeting vector. **A:** Step 1: A minimal vector containing an ampicillin-resistance cassette, and a DTA sequence for negative selection, was digested into a linearized construct. A PCR construct was generated, using oligos with 50 bp long arms (purple) that were homologous to regions in the modified MYL2 BAC, 8.8kb upstream of the modified region (Oligo_F), and 2.3kb downstream (Oligo_R). Step 2: The MYL2 modified region in the BAC was subcloned into the PCR construct through the homologous regions between BAC and PCR construct ends (purple). **B:** A PCR was performed to screen for correctly recombineered subclones. **C:** Integrity of a selection of subclones was checked by digestion with either BglI or XbaI. Clone 3.9 was used for further MYL2 genomic targeting in hESCs.

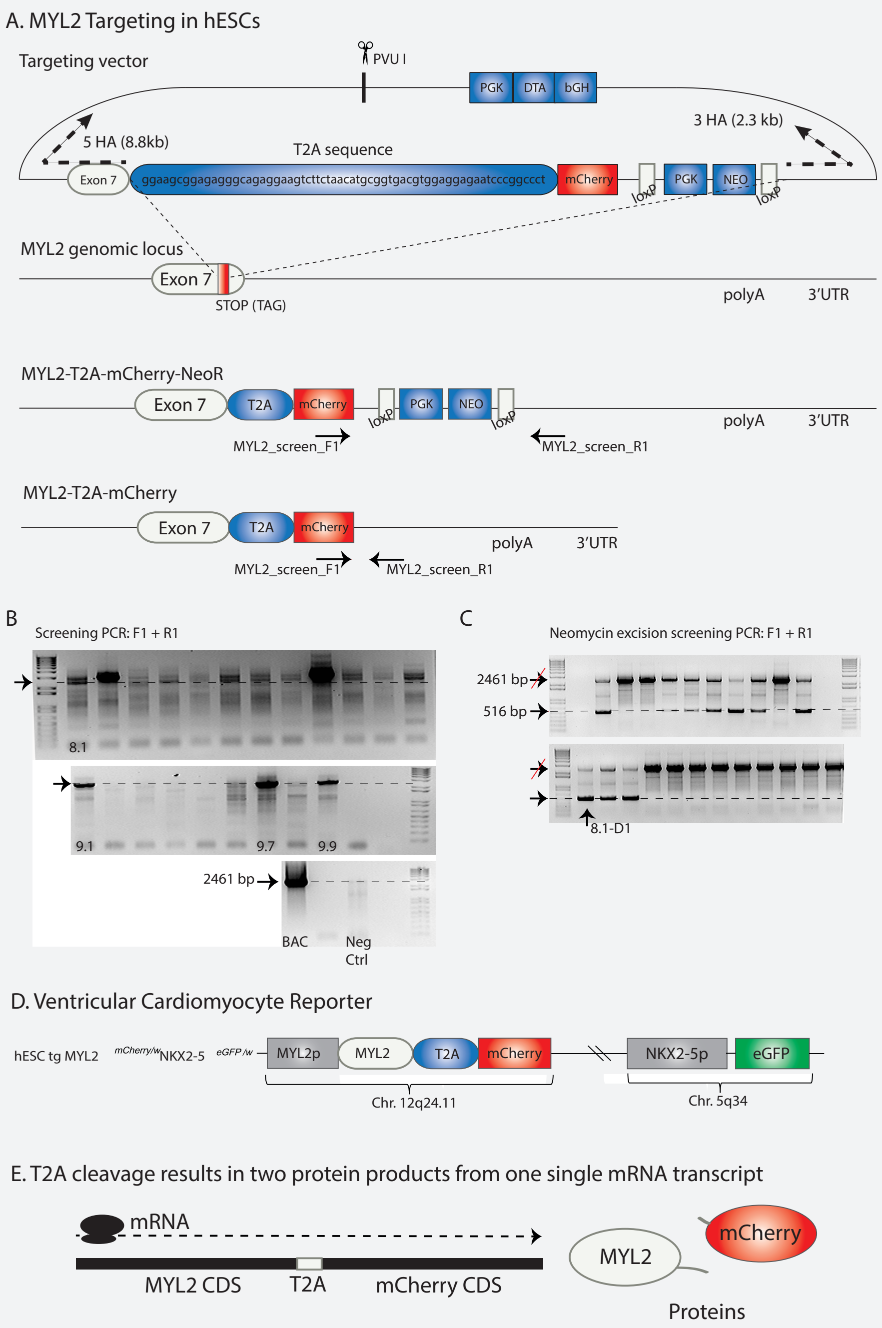


Figure 3. MYL2 targeting in NKX2-5$^{eGFP/w}$ hESCs. A: The MYL2 targeting construct was designed to replace the MYL2 stopcodon, located in exon 7. Prior to targeting, the construct was linearized through PvuI digestion, and electroporated into NKX2-5eGFP/w hESCs[9]. Targeted cells were selected through their resistance to G418, which was added to the culture medium after electroporation (Oligos used: MYL2_screen_F1 + MYL2_screen_R1). Cre-mediated excision of the selection cassette (through the presence of two distal loxP sites) resulted in the final MYL2-T2A-mCherry reporter line. **B:** PCR screening of selected clones. Positive control was the modified MYL2 BAC, showing a 2461 bp band. Negative control was the unmodified NKX2-5$^{eGFP/w}$ hESCs line. **C:** PCR screening on neomycin cassette excision, which would result in a 516 bp band, and the absence of the 2461 bp band. **D:** Overview of the finally generated dual cardiac reporter line. **E:** Schematic drawing of the ribosomal translation of one mRNA transcript containing both MYL2 CDS and mCherry CDS. Cleavage of the T2A sequence during ribosomal translation results in two separate proteins.

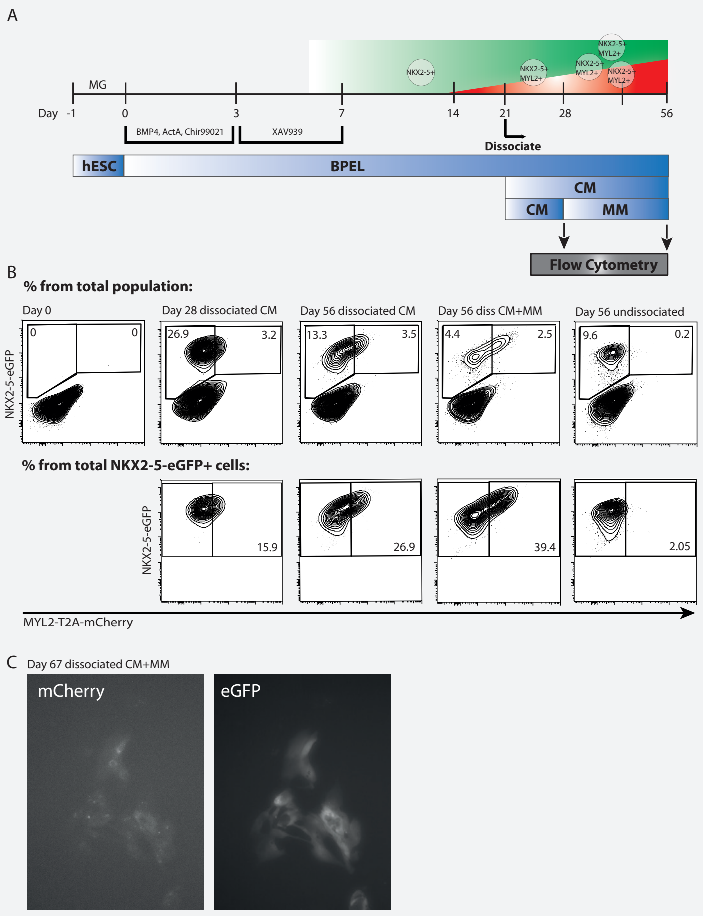


Figure 4. Directing hESC differentiation towards ventricular cardiomyocytes. A: Overview of cardiac monolayer differentiation protocol. hESCs were plated on matrigel, one day prior to induction with BPEL containing BMP4, Activin-A, and GSK3-inhibitor Chir99021. After three days, growth factors were replaced with BPEL containing Wnt signalling inhibitor Xav939. This was replaced with plain BPEL for further continuous culture. At day 21 of differentiation, cells were dissociated and further cultured.

BAC Quadruple Recombineering	
Oligo1	TGGACTACAAGAACCTGGTGCACATCATCACCCACGGAGAAGAGAAGGACggaagcggagagggcagaggaagtcttctaacatgcggtg
Oligo2	CAGGGACCACTCTGCAAAGACGAGCCCAGGGCGCAGCAGCGAGCCCCCTCGATGGATGCAGAGTGTGCGTAGACTATGAACGTACTTAAC
BAC screening	
SV40-screen_F	CCCCCTGAACCTGAAACATA
MYL2_screen_R1	GCAAAGAAGATGGAGGTGGA
Subcloning	
Oligo_F1	CTGTCAAGTGGATACTATTATTATCATCCCCATTTTACAGGAGAGGCAACGTCGACGCTCTCCTGAGTAGGACAAATC
Oligo_R1	cagagttggtccaggccgagagcctcatacccgggctcctgtggtcagagtcgacTCCGCCTCAGAAGCCATAGA
Subcloning screening	
SV40-screen_F	CCCCCTGAACCTGAAACATA
mmvector-DTA-seq_F1	CGATCTCTTTGTGAAGGAACC
MYL2 Targeting Screening	
MYL2_screen_F1	CCCCGTAATGCAGAAGAAGA
MYL2_screen_R1	GCAAAGAAGATGGAGGTGGA
MYL2 Neomycin Excision Screening	
MYL2_screen_F1	CCCCGTAATGCAGAAGAAGA
MYL2_screen_R1	GCAAAGAAGATGGAGGTGGA

Table 1. Oligo sequences

Genes	Forward primer	Reverse primer
hARP	CACCATTGAAATCCTGAGTGATGT	TGACCAGCCCAAAGGAGAAG
MYL2	GATGTTCGCCGCCTTCCCCC	GCAGCGAGCCCCCTCCTAGT
MYH6	GCTGGCCCTTCAACTACAGA	CTTCTCCACCTTAGCCCTGG
MYH7	GAGGACAAGGTCAACACCCT	CGCACCTTCTTCTCTTGCTC

Table 2. Quantitative PCR primers for human genes

REFERENCES

1. Spater D, Hansson EM, Zangi L, et al. How to make a cardiomyocyte. DEVELOPMENT 2014;141(23):4418–4431.
2. Bruneau BG. The developmental genetics of congenital heart disease. NATURE 2008;451(7181):943–948.
3. Matsa E, Burridge PW, Wu JC. Human Stem Cells for Modeling Heart Disease and for Drug Discovery. SCIENCE TRANSLATIONAL MEDICINE 2014;6(239):239ps6–239ps6.
4. Laslett LJ, Alagona P, Clark BA, et al. The worldwide environment of cardiovascular disease: prevalence, diagnosis, therapy, and policy issues: a report from the American College of Cardiology. J. AM. COLL. CARDIOL. 2012;60(25 Suppl):S1–49.
5. Passier R, van Laake LW, Mummery CL. Stem-cell-based therapy and lessons from the heart. NATURE 2008;453(7193):322–329.
6. Burridge PW, Matsa E, Shukla P, et al. Chemically defined generation of human cardiomyocytes. NAT METH 2014:–.
7. Blazeski A, Zhu R, Hunter DW, et al. Cardiomyocytes derived from human induced pluripotent stem cells as models for normal and diseased cardiac electrophysiology and contractility. PROGRESS IN BIOPHYSICS AND MOLECULAR BIOLOGY 2012;110(2-3):166–177.
8. Hartogh Den SC, Schreurs C, Monshouwer-Kloots JJ, et al. Dual Reporter MESP1 mCherry/w-NKX2-5 eGFP/whESCs Enable Studying Early Human Cardiac Differentiation. STEM CELLS 2014;33(1):56–67.
9. Elliott DA, Braam SR, Koutsis K, et al. NKX2-5eGFP/w hESCs for isolation of human cardiac progenitors and cardiomyocytes. NAT METH 2011;8(12):1037–1040.
10. Chuva de Sousa Lopes SM, Hassink RJ, Feijen A, et al. Patterning the heart, a template for human cardiomyocyte development. DEV. DYN. 2006;235(7):1994–2002.
11. MULLER M. Selection of ventricular-like cardiomyocytes from ES cells in vitro. THE FASEB JOURNAL 2000;14(15):2540–2548.
12. Lee MY, Sun B, Schliffke S, et al. Derivation of functional ventricular cardiomyocytes using endogenous promoter sequence from murine embryonic stem cells. STEM CELL RESEARCH 2012;8(1):49–57.
13. Trichas G, Begbie J, Srinivas S. Use of the viral 2A peptide for bicistronic expression in transgenic mice. BMC BIOLOGY 2008;6(1):40.
14. Donnelly ML, Gani D, Flint M, et al. The cleavage activities of aphthovirus and cardiovirus 2A proteins. J. GEN. VIROL. 1997;78 (Pt 1):13–21.
15. Kim JH, Lee S-R, Li L-H, et al. High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. PLOS ONE 2011;6(4):e18556.
16. Costa M, Dottori M, Sourris K, et al. A method for genetic modification of human embryonic stem cells using electroporation. NATURE PROTOCOLS 2007.
17. Chan HY, V S, Xing X, et al. Comparison of IRES and F2A-based locus-specific multicistronic expression in stable mouse lines. PLOS ONE 2011;6(12):e28885.
18. Ha S-H, Liang YS, Jung H, et al. Application of two bicistronic systems involving 2A and IRES sequences to the biosynthesis of carotenoids in rice endosperm. PLANT BIOTECHNOL. J. 2010;8(8):928–938.
19. Provost E, Rhee J, Leach SD. Viral 2A peptides allow expression of multiple proteins from a single ORF in transgenic zebrafish embryos. GENESIS 2007;45(10):625–629.
20. Mummery CL, Zhang J, Ng ES, et al. Differentiation of Human Embryonic Stem Cells and Induced Pluripotent Stem Cells to Cardiomyocytes A Methods Overview. 2012.
21. Huber I, Itzhaki I, Caspi O, et al. Identification and selection of cardiomyocytes during human embryonic stem cell differentiation. THE FASEB JOURNAL 2007;21(10):2551–2563.
22. Klein I, Ojamaa K. Thyroid Hormone and the Cardiovascular System. N ENGL J MED 2001;344(7):501–509.
23. 23.Yang X, Pabon L, Murry CE. Engineering adolescence: maturation of human pluripotent stem cell-derived cardiomyocytes. CIRCULATION RESEARCH 2014;114(3):511–523.
24. Krüger M, Sachse C, Zimmermann WH, et al. Thyroid hormone regulates developmental titin isoform transitions via the phosphatidylinositol-3-kinase/ AKT pathway. CIRCULATION RESEARCH 2008;102(4):439–447.

CHAPTER 7

Analysis of Pre-Cardiac MESP1 Progenitor Proliferation at Single Cell Level using an eGFP-Anillin Reporter Line

Sabine C. den Hartogh[1], Lu Cao[1], Tom Zwetsloot[1], Alexandra Raulf[2], Aniek van Wijngaarden[1], Michael Hesse[2], Bernd K. Fleischmann[2], Christine Mummery[1], Robert Passier[1,3].

1. Leiden University Medical Center, Dept. Anatomy and Embryology, Leiden, The Netherlands.
2. Institute of Physiology I, Life and Brain Center, University of Bonn, 53105 Bonn, Germany.
3. Department of Applied Stem cell Technologies. MIRA Institute for Biomedical Technology and Technical Medicine. University of Twente, P.O.Box 217, Enschede, The Netherlands.

ABSTRACT

Production of cardiomyocytes from human pluripotent stem cells (PSCs) provides not only promising opportunities for both *in vitro* as well as *in vivo* translational applications, but also may lead to a better understanding of signalling pathways and biological processes during the early phases of cardiac differentiation and development. For translational applications in drug development, disease modelling and cell-based therapy, it is important to obtain a large amount of highly purified cardiomyocytes from PSCs. Alternatively, controlled and defined expansion of cardiac progenitor cells offers additional advantages. Here, we studied the proliferative capacity of pre-cardiac MESP1 progenitors through the combinatorial use of fluorescent proliferation reporter eGFP-Anillin and the previously generated MESP1-mCherry cardiac reporter line. Scaffolding protein anillin is localized at distinct locations during the cell cycle, and fusion to eGFP enables distinction from cytokinesis with other cell cycle stages. It is currently unknown what the proliferative capacity is of MESP-positive pre-cardiac progenitors and how they respond to changes in environmental conditions. Therefore, we studied the proliferative capacity by analysis of subcellular localization of eGFP-Anillin in MESP1-mCherry progenitors under controlled culture conditions and using time-lapse microscopy. In addition, we screened several small molecules and growth factors on the ability to enhance MESP1 progenitor proliferation.

INTRODUCTION

Heart disease is the most common cause of death in most Western societies, unlocking the potential of regenerative stem cell- and drug therapy[1]. For successful development of cell-based therapies, large numbers of human cardiac cells will be necessary. Human pluripotent stem cells (hPSCs) can be efficiently differentiated to cardiac cells[2, 3], however, the inverse relationship between the degree of differentiation and the proliferation rate of cardiomyocytes limits the production of cardiac cells and advocates striving for an expandable cardiac progenitor population. When hPSCs, including induced PSCs (hiPSCs) and embryonic SCs (hESCs), differentiate towards the cardiac cells they first transit from a mesendodermal stage, expressing genes such as Brachyury T, MIXL1, and Eomesodermin, to a pre-cardiac mesodermal stage, marked by the expression of transcription factor MESP1[4]. These cells then further segregate to distinct multi-potent cardiac progenitor cells (CPCs), and start to express cardiac transcription factor NKX2-5[5]. These CPCs are capable of further differentiation into several cardiac cell subtypes, including atrial, ventricular, pacemaker-like cardiomyocytes, but also endothelial, and smooth muscle cells. In order to obtain high yields of cardiac subtype populations and to control maintenance and expansion of CPCs, it is of interest to understand the underlying mechanisms[6-9]. Several studies have demonstrated a crucial role for FGF-, IGF-, Notch1, and Wnt/B-catenin signaling pathways for self-renewal/expansion of CPCs. Nevertheless, the self-renewing capacity of specific CPCs and their biological responses may vary between different CPCs and/or developmental stages. As we know from *in vivo* studies in mice, almost all cells of the heart are derived from MESP1 expressing progenitors[10]. Therefore, it is for both regenerative and drug development studies of interest to understand whether we are able to maintain and expand human MESP1 progenitors in culture.

In order to study the proliferative capacity of MESP1 progenitors, we made use of a *in vivo* reporter system using the scaffolding protein anillin fused to enhanced green fluorescent protein (eGFP), to provide high spatiotemporal resolution of the mitotic phase[11]. Anillin is localized at specific subcellular locations in the cell during cytokinesis and midbody formation, enabling a clear visual separation of the different phases of the cell cycle, including cell division (**Fig 1**). Anillin is localized in the cell nucleus during G1-phase, S-phase, and G2-phase, and moves to the cytoplasm when entering the M-phase. Prior to cell division, anillin locates in the cytoskeleton ring and becomes restricted to the midbody region upon cytokinesis. When cells leave an active cell cycle, anillin gets degraded.

Here, we generated a dual fluorescent reporter line by introducing eGFP-Anillin in the MESP1$^{w/mCherry}$ hESC line, allowing live imaging of the different phases of cell division of cardiac progenitors and their derivatives[11]. Using time-lapse microscopy we were able to visualize the dynamics of MESP1-expressing progenitors. We measured a low proliferative capacity of these progenitors during *in vitro* cardiac differentiation. Further, we optimized this model for screening cytokines/small molecules for their ability to enhance MESP1-progenitor proliferation. We further discuss future development of automated analysis of cell proliferation, using the eGFP-Anillin reporter, and we indicate its potential in the development of cardiomyocyte-proliferation assays, which will be of interest for the fields of regenerative medicine and assay development for drug discovery and toxicity[12-15].

METHODS

Lentiviral transductions in hESCs

To generate the hESC proliferation reporter line, 100k MESP1$^{w/mCherry}$ hESCs[4] were transfected with a lentiviral construct containing a CAG–eGFP–anillin expression cassette (0.5uL / 20k cells), and plated on a Matrigel (BD Biosciences; growth factor reduced, phenol red–free)-coated 24 wells plate (Falcon)[11]. Three days after transduction, hESCs were enzymatically passaged using 1x TrypLE Select (Life Technologies). After dissociation, cells were stained with anti-SSEA-4 in order to select for pluripotent hESCs, followed by resuspension in FACS sorting buffer (PBS with 0.5% BSA and 2.5mM EDTA) and filtration through a 40-µm cell strainer (Falcon). Cells were sorted using a BD ARIA III flow cytometer into 96 wells flat bottom low attachment plates. Live cells were gated on the basis of side scatter, forward scatter, anti-SSEA-4-PE, and eGFP-Anillin fluorescency. Flow cytometric gates were set using control cells labelled with the appropriate isotype control antibody.

Clone Selection and Culture

Due to the integration properties of lentiviral particles, five undifferentiated transfected clones were analysed for their eGFP-Anillin expression. Clone A4 and C6 were selected, based on their eGFP expression intensity, and further cultured as single cells on mouse embryonic fibroblasts on 9.96 cm2 6 well format (Falcon), in 3 mL hESC medium (DMEM F12, Non Essential Amino Acids, Knock Out Serum (Gibco by Life Technologies); bFGF (Miltenyi Biotec)) and were enzymatically passaged. hESCs from clones A4 and C6 were analysed for eGFP-Anillin expression. Clone A4 showed a more consistent and robust eGFP-Anillin expression and further experiments were performed with this clone.

Cardiac Differentiation

One day before differentiation, 10k hESCs were passaged onto a 96-well tissue culture dish (96 Well Optical CVG, ThermoScientific), coated with Matrigel (BD Biosciences) in 100uL of hESC medium per well. On the first day of differentiation, day 0, hESC medium was replaced by low insulin serum-free medium BPEL (BSA, polyvinyl alcohol, essential lipids, as previously described)[17, 24], containing the cytokines BMP4 (20 ng ml−1, R&D Systems) and Activin A (20 ngml−1, Miltenyi Biotec), and Chir99021 (1.5uM, AxonMedchem). At day 3 of differentiation, the medium was refreshed with BPEL containing 5uM Xav939 (R&D). To study their proliferation, we sorted MESP1-mCherry progenitors at day 2 of differentiation. Cells were plated back as monolayers in 96 wells plates (96 Well Optical CVG, ThermoScientific), in a density of 50k cells per well, in BPEL with 5uM Xav939 from day 2 until day 7.

Growth Factor Screening

For the *in vitro* growth factor screening, cells were differentiated as described before, on 96 well tissue culture dishes (Falcon). A complete list of the compounds used for screening, and the concentrations used, can be found in **Supplemental Table 1**. Growth factors were added to the cell culture at day 2 or day 3 of differentiation. The proliferative effect was assayed on day 4 of differentiation using a MACS Quant VYB flow cytometer (Miltenyi Biotec). The percentage of MESP1-mCherry+ cells was assessed for each treatment, and compared to control cells without treatment. Each experiment was performed in quadruplicate, and one or two independent experiments were performed for each condition.

Flow Cytometry

During cardiac differentiation, eGFP-Anillin and MESP1-mCherry expression were analysed every 24 hours. For this, cells were enzymatically dissociated using 1x TrypLE Select (Life Technologies). After dissociation, cells were resuspended in FACS sorting buffer (PBS with 0.5% BSA and 2.5mM EDTA) and analysed using a MACS Quant VYB flow cytometer (Miltenyi Biotec) and FlowJo Software with a 488 nm and a 561 nm laser. Co-expression of eGFP-Anillin with proliferation marker Ki-67 was analysed by flow cytometry in undifferentiated hESCs, and at differentiation day 3 and day 5. For this, dissociated cells were treated with FIX & PERM® Cell Fixation and Permeabilization Kit (Life Technologies) and stained with anti-Ki-67-PE (Biolegend).

Immunohistochemistry

eGFP-Anillin transduced MESP1-mCherry-cells were analyzed by immunohistochemistry (IHC) when undifferentiated, and at day 3, and 5 of differentiation. For this, cells were fixed in 2% PFA and stained for Ki-67 (MIB-1 M7240, DAKO (1:400)), PHH3 (Ser 10, Millipore 06-570 (1:100)), and aurora B kinase (SIGMA A5102 (1:200)). Primary antibodies were diluted in 0.2% TritonX in PBS, supplemented with 5% Donkey serum and incubated with the cells for 1.5 hour at room temperature (RT). Then, cells were washed 3x 5 minutes with PBS. The 2nd antibodies (mouse and rabbit Cy5 Jackson Immuno Research) were directly diluted (1:400) in Hoechst 33342 solution (1 µg/ml) in PBS and incubated for 1 hour at RT in dark. Then, cells were washed 3x with PBS and embedded in FLUKA mounting medium (Sigma10981) on glass-cover slips.

Statistical Methods

A two-tailed Student's T test was performed using Graphpad Prism software (Graphpad Software, La Jolla, CA, USA). Results were considered significant at P values <0.05.

Time Lapse Microscopy

In order to study cellular dynamics of MESP1 progenitors, eGFP-Anillin-MESP1$^{w/mCherry}$ hESCs were differentiated as previously described, and sorted at day 3 of differentiation. 50k cells were replated as monolayer onto a 96-well glass bottom plate (Thermo Scientific), coated with Matrigel. The cells were incubated on a stage-heated fluorescence image acquisition station (Leica AF6000). The plate was covered with a CO2 environment cover glass, supplied with 8% CO2/O2 inflow to maintain pH balance. Time-lapse imaging was started 5 hours after plating. Imaging settings can be found in supplemental table 2 (**Supp. Table 2**). Cell tracking was performed manually using the Fiji ImageJ plugin MTrackJ (E. Meijering, Methods in Enzymology 2012).

Automated Quantification of Proliferation

The pipeline of automated quantification of eGFP-Anillin localizations at single cell level contains two major parts: image analysis and data analysis. Image analysis procedure basically converts raw microscope images into quantified measurements representing characteristic biological phenomena. Three steps are elaborated to achieve this purpose: (1) image enhancement, (2) image segmentation and (3) phenotype measurement. Here, image enhancement aims at providing a better input for image segmentation by a list of techniques including background correction, noise suppression and contrast enhancement. Image segmentation partitions an image into multiple regions (single cell regions) with the goal to simplify the representation of an image into comprehensive components. For fluorescence microscopy cell imaging, we utilized seeded propagation cell segmentation method[16]. The method segments the single nuclei using watershed segmentation based on Euclidian distance map (EDM) from the DAPI channel. The segmented single nuclei mask is treated as seed and propagate to find the Anillin positive cell area in the GFP channel. Phenotype measurements including shape and texture features are calculated on both individual nuclei and Anillin positive cell area masks. These phenotype measurements are further used as feature dataset in the data analysis part. In order to quantify the number of dividing cells versus non-dividing cells, a classifier was trained to distinguish the dividing cells from the image. Supervised classification strategy, which is based on ground truth analysis (dividing = 1 versus non-dividing = 0), was used in this study. First, a feature dataset was normalized in order to equalize the ranges of the features.

Second, the feature selection procedure was applied to select a subset of relevant features. Four representative feature selection methods were evaluated: the best individual-N features, the branch and bound procedure, greedy forward selection and greedy backward elimination[17]. Third, four classifiers covering both linear and non-linear categories; i.e. the linear classifier (LDC), the quadratic classifier (QDC), k-nearest neighbour classifier (KNNC) and support vector machine (SVC) were included for the classification evaluation. The final performance of the automated analysis was evaluated using the precision, recall and F-score[16, 18].

RESULTS

eGFP–Anillin expression marks proliferating hESCs.

In order to study the proliferative capacity of early pre-cardiac MESP1-mCherry progenitors, we transduced previously generated MESP1$^{w/mCherry}$ hESCs[4] with lentiviral particles containing a CAG-eGFP-Anillin fusion cassette[11], resulting in random genomic integration of the CAG-eGFP-Anillin reporter in actively transcribed genes (**Fig. 2**). Five targeted clones were obtained (showing the brightest eGFP expression in undifferentiated hESCs), from which clone A4 showed the most consistent expression (**Fig 3a**). In clone A4, eGFP-Anillin expression was analyzed for its co-expression with M-phase marker PHH3, proliferation marker Ki-67 (S, G1, G2, M phase), and aurora B (midbody expression) (**Fig 3b,c**). Indeed, we found correlating expression patterns of eGFP-Anillin with PHH3, aurora B co-expression with eGFP-Anillin in the midbody of dividing cells (telophase), and all cells positive for both eGFP-Anillin and Ki-67, indicating how different eGFP-Anillin localizations can distinct specific cell cycle stages from eachother.

eGFP-Anillin expression intensity decreases upon cardiac differentiation

To further characterize the MESP1$^{w/mCherry}$ eGFP-Anillin reporter, we differentiated cells (clone A4) as described before. We analyzed eGFP and mCherry expression every 24 hours from day 0 until day 5 of differentiation (**Fig. 4**). We found MESP1-mCherry levels peaking around day 2 and day 3 (12-14%). Furthermore, we observed all cells being positive for eGFP when in undifferentiated state, up to day 2 of differentiation. However, we found a rapid decrease in eGFP-Anillin intensity upon day 3 of differentiation, where we could only detect 75% of cells still positive for eGFP-Anillin. These levels decreased to only 37% at day 5 of differentiation, indicating cells entering the G0-phase upon differentiation. An alternative explanation for decreased expression of GFP can also attributed to epigenetic silencing of the transgene (for example by DNA-methylation and chromatin remodelling). To further investigate this, we performed analysis on Ki-67 co-expression at day 3 and day 5 of differentiation and compared this to eGFP-Anillin (**Fig. 5a,b**). Despite decreasing levels of eGFP-Anillin, we could still detect maintained expression of Ki-67 in a large population of cells (±80% at day3, and ±77% at day 5), indicating silencing of the construct upon differentiation at day 5. Moreover, at day 5 of differentiation, we could identify cells positively stained for Ki-67, and negative for Anillin-eGFP (**Fig. 5b**). The same image shows high variability of eGFP levels between cells, at both day 3 and 5 of differentiation. Based on these results we concluded that from day 5 onwards silencing of the transgene would interfere with the experimental outcome. Therefore, time-lapse imaging and analysis was performed up to day 4 of cardiomyocyte differentiation.

Proliferation of Early Cardiac Progenitors

In order to get insight in the proliferative capacity of early cardiac progenitors *in vitro*, we manually analysed cell division in the first four days of differentiation (**Fig. 6a**). Based the number of cells that was expressing eGFP-Anillin in the cytoplasm, cytoplasmic ring, and midbody region, the percentage of proliferating hESCs was measured to be approximately 4%, which was similar to the findings observed by Hesse et al.[11]. This proliferative capacity was largely decreased upon the first days of differentiation, and slightly increased after the formation of mesoderm (day 3,4). MESP1 progenitor division could be nicely visualized through the co-expression of mCherry with eGFP-Anillin localizations in the cytoplasm, cytoskeleton, and midbody. Moreover, dividing cells became rounded and were less spread out. They showed to have a proliferative capacity of about 2-3%, with no significant difference between early (day 2) and late (day 4) progenitors (**Fig 6a,b**). In order to get more insight in the proliferative capacity of MESP1-mCherry progenitors at single cell level, we performed time lapse imaging during cardiac monolayer differentiations, and of sorted MESP1-mCherry progenitors.

Induction of MESP1-mCherry proliferation

To identify signalling pathways involved in early cardiac progenitor proliferation, MESP1-mCherry–eGFP-Anillin hESCs were differentiated as described before and treated with growth factors/small molecules on day 2 or day 3 of differentiation. As a preliminary screen we measured the yield of MESP1-mCherry+ cells by automated flow cytometry, at day 4 of differentiation (**Fig 7**). The control treatment did not contain any GFs (plain BPEL). For all growth factors we have added in our screening, we could not find a significant upregulation of mCherry+ cells, compared to control, indicating no aberrant role for these pathways in MESP1-progenitor proliferation under the tested conditions, which is in contrast to their role in proliferation of other early cardiac progenitors[6, 9, 19]. Therefore, we postulate that a complex combination and concentration of growth factors and possibly specific ECM components, which are highly expressed by MESP1-mCherry progenitors[4], may be required.

Real time tracking of MESP1-mCherry progenitors

In order to study the proliferative capacity of CPCs at single cell level, we developed a live single cell tracking system during differentiating, to follow division and migration of single cell MESP-mCherry expressing progenitors. For this, optimization of culture conditions and technical settings were required, which included light intensity of the fluorescent lamp, exposure time, autofocusing, heating temperature, objective magnification and objective immersion medium (**Supp. Table 2**). To decrease the effects of photo-bleaching, we used a low lamp intensity (4.3) and 2x2 binning (**Fig. 8a**). Another setting that required optimization included Z-position stability, or focal drift. Without autofocus, focal drift would occur, which makes long time-lapse experiment impossible. Since eGFP-Anillin signal diminishes easily, autofocus was set up in the bright field-DIC channel. This resulted in a consistent error in the focal plane of 5-6 μM, which was compensated by making images in multiple Z-planes.

In order to generate an assay to study the proliferative capacity of MESP1 progenitors live, we sorted MESP1-mcherry progenitors at day 3 of differentiation and replated them as monolayers at 96 wells plates with glass bottom. We started live-image acquisition after cell attachment (about 5 hours later). We used a maximum interval time of 10 minutes to be able to track single cells (**Fig. 8b**). Live cell division could be followed through the distinct localizations of the eGFP-Anillin reporter protein (**Fig. 8c**). We identified how the nucleus changes from oval to a more U-shaped form prior to cell division, and how cell shape became more rounded when eGFP-Anillin moved to the cytoplasm during the prophase. These changes prior to cell division are described as mitotic cell rounding. This was followed by the formation of a cytoplasmic ring, and a midbody upon cytokinesis. We estimated that the M-phase of early cardiac progenitors would take about 30 minutes, which is consistent with previous findings from Hesse et al. who demonstrated a similar time for the M-phase of hPSCs[11]. In a similar sort-and-replate experiment, we determined the migrating behaviour of MESP1-mCherry expressing progenitors, using a time interval of 15 minutes (**Fig. 8d**). We were able to follow MESP1-mCherry progenitors for up to 15 hours after replating and attachment (20 hours after sorting). Some tracked cells were lost, as they migrated out of focus. Further, we showed the ability to visualize MESP1-mCherry progenitor proliferation, and identified the distinct eGFP-anillin localizations marking the different stages of the M-phase and final cytokinesis (**Fig. 8e**).

To understand the dynamics of MESP1-mCherry progenitor proliferation and their further differentiation towards cardiomyocytes, it will be of interest to follow their behaviour and phenotype changes in time, through time lapse imaging at single cell level. Here, we show how sorted and replated MESP1-mCherry progenitors, were further differentiating over time, followed for up to 48 hours (**Fig. 9**). Bright MESP1 expression at day 3 was diminished upon further differentiation, and cell shape and behaviour changes could be observed. eGFP-anillin signal faded alongside MESP1-mCherry signal and was completely gone at day 5 of differentiation, possibly due to silencing upon differentiation (**Fig. 9**).

Automated Quantification of Proliferation

In order to study whether small molecules/growth factors may enhance proliferation of cardiac progenitors and/or cardiomyocytes using the eGFP-Anillin cassette, we developed a pipeline for automated quantification of eGFP-Anillin localizations at single cell level. For the image analysis part, we used a seeded propagation cell segmentation method to identify the single cells from the image. The precision of seeds (single nucleus) picking is crucial for a correct propagation procedure. The EDM watershed segmentation method works best for smooth convex objects that don't overlap too much. However, in our case, when cells are dividing, the shape of nuclei becomes irregular and dispersed. Furthermore, highly overlapping nuclei regions were observed in our image dataset. Therefore, we introduced a merging criterion which is based on the homogeneity of region intensity for the over segmented regions of nuclei under division. Since the number of highly overlapped nucleus regions is limited, these regions were excluded for the phenotype measurement. An example of the final segmentation result is shown in **Figure 10**. For the data analysis, 1500 individual cells (291 dividing cells and 1209 non-dividing cells) and 44 features were firstly collected as training dataset. During the classifier training, each combination of a feature selection method and a classifier algorithm was repeated for 100 times and the performance for each combination was evaluated using weighted classification error, which assigns different importance to different classes (www.mathworks.com)[20].

Based on this, we found that the combination of branch and bound feature selection (B&B) with k-nearest neighbour classifier (KNNC) has the lowest minimal value of mean weighted error and relatively small standard deviation of mean, as can be concluded from **Table 1**. The minimal value was reached when the feature selection method selects 4 features. These 4 features include smoothness of cell intensity, area of eGFP-Anillin positive cell area region, entropy of eGFP-Anillin signal in nuclei region, standard deviation of nuclei intensity.

Next, to identify the degree of errors in the automated analysis, we compared the results from manually quantifications of 22 images with their automated quantified counterpart as shown in **Figure 11a-d**. To demonstrate quantitatively the accuracy of the automated classification, we have used the Precision, Recall, and F-score metrics[21] **(Fig. 11e)**. We reached a F-score equalling ±0.85, which meets or exceeds the performance of previous cell segmentation and/or pattern recognition studies[16, 22].

DISCUSSION

The availability of large numbers of cardiac committed progenitor populations is of high interest for translational purposes, such as cell therapy or drug discovery. An additional advantage is that it may provide us with a model for studying cardiac progenitor cells in relation to heart development. Although numerous studies have been conducted to unravel molecular and morphological aspects of heart development, in-depth knowledge regarding formation, proliferation, (subtype) differentiation and self-renewal of cardiac progenitors is limited. From *in vivo* mouse embryo studies it is evident that during gastrulation, cells ingress through the primitive streak and start to express mesendodermal transcription factors, including Brachyury T and Eomesodermin. From these cells, a posterior subpopulation develops that expresses key cardiac transcription factor MESP1. Lineage tracing studies in mouse have shown that the vast majority of the cells of the heart are derived from this MESP1-expressing population. In order to study the characteristics of these early pre-cardiac progenitor population in human, we previously developed a dual cardiac reporter MESP1$^{mCherry/w}$NKX-2-5$^{eGFP/w}$ hESC line[4]. This allowed us to isolate MESP1-mCherry expressing progenitors and to follow their further differentiation towards NKX2-5-eGFP expressing cardiac cells. In our current study, we were interested in the proliferative capacity of these MESP1 progenitors, in order to understand their potential as an unlimited source of cardiac progenitors. Several methods for measuring cell proliferation exist, on fixed and living material. M-phase marker PHH3 and active cell cycle marker Ki-67 are commonly used for measuring proliferation at fixed time-points. However, both markers require fixation and staining of samples, and are not appropriate for quantifying live cell division. Moreover, they are not specifically marking cytokinesis, and therefore do not distinguish between endoreduplication and cell division[11]. To overcome these limitations we made use of an *in vitro* reporter system that was recently developed[11], using the scaffolding protein anillin fused to eGFP. Here, we transduced MESP1$^{mCherry/w}$NKX-2-5$^{eGFP/w}$ hESCs with lentiviral particles containing the CAG-eGFP-Anillin cassette, and thereby developed an *in vitro* system to study the proliferation of early cardiac progenitors, up to day 5 of differentiation. Since activation of NKX2-5-eGFP occurs at later stages during differentiation (~day 7 and later), GFP expression at early stages could be attributed to Anillin-eGFP alone. In hESCs, we showed the different stages of the M-phase by distinct localizations of eGFP-Anillin, clearly discriminating cytokinesis from other cell cycle stages. We showed that the vast majority of eGFP-Anillin+ hESC co-localized with Ki-67 immunostaining, whereas dividing cells co-localized with PHH3. Upon differentiation of hESCs towards the cardiac lineage, we found decreasing eGFP levels, leading to absence of detectable GFP levels at day 6. One possible explanation for this decrease in GFP expression could be due to an increasing number of cells entering a quiescent cell cycle stage upon differentiation[23], although this is not very likely, since we only observed a slight decrease in Ki-67 levels during the first days of differentiation. A more plausible explanation is silencing of the gene in which the CAG-eGFP-Anillin cassette had integrated[24]. As the Anillin-eGFP signal could still be detected up to day 5 of differentiation, the generated reporter line allowed us to study early progenitor differentiation. By measuring at fixed timepoints, we found, that approximately 4% of undifferentiated hESCs were dividing. MESP1 progenitors show a similar percentage, with no difference between early (day 2) and late (day 4) MESP1 progenitors. To fully understand the proliferative capacity of MESP1 progenitors at single cell level, and to get insights into the dynamics of these progenitors prior their further differentiation towards cardiomyocytes, we decided to set up a time-lapse live imaging system. Many hurdles needed to be overcome, including photo bleaching of the fluorescent protein upon exposure to high light intensities, the ability to track migrating single cells (time interval of image acquisition, autofocus, and Z-stacks), and the amount of data generated from a long-term time-lapse study (3-5 days, 10 minute interval, 3 Z-stacks). Following optimization by adjusting these parameters we showed real-time tracking of MESP1-mCherry progenitors during proliferation and further differentiation. In future studies we will follow differentiation from day 0 up to day 5, where we will study the derivation of MESP1-mCherry expressing progenitors, and quantify their proliferative behaviour prior to their further differentiation towards cardiomyocytes. In addition, time lapse experiments at later stages of differentiation, in the previously developed MESP1$^{mCherry/w}$NKX-2-5$^{eGFP/w}$ hESC line, will be of interest to perform for a more complete understanding of the dynamics of MESP1 expressing progenitors and their further differentiation towards specific cardiac lineages. Moreover, time-lapse experiments on cell proliferation may also enlighten knowledge on circadian rythms of cardiac progenitors, for which evidence grows it is tightly regulating cellular function, such as proliferation and/or maturation[25].

For potential future high-throughput screening experiments on proliferation of cardiac progenitors or cardiomyocytes, we are currently working on the development of an automated quantification method using classifiers that are based on the distinct localizations of eGFP-Anillin. Using such an automated system, a large number of combinations of growth factors and small molecules can be screened for their enhancing effect on MESP1-progenitor proliferation. Our automated image analysis system is efficient to identify single cells from the image. However, segmentation of the nucleus with diverse morphology or in a highly overlapped region needs further improvement. Nevertheless, after selecting the most optimal automated classification method for our set-up, we managed to reach a high F-score equalling ±0.85, which exceeds the performance of previous cell segmentation studies. In further experiments, we will test our newly developed quantification method in the absence and presence of proliferation inhibitors and stimulators, to understand which improvements are desired for implementation of this method in large-scale screening experiments. Automated image analysis is a critical step for the development of phenotypic screens that allow high-speed, reproducible and quantitative analysis.

In order to identify crucial signalling pathways that play a role in cardiac progenitor proliferation, which would subsequently be of interest as target pathways in large-scale screening experiments, pilot experiment were performed that included initial screenings on our standard monolayer differentiations. At day 2 or day 3 of differentiation, we removed the standard cardiac differentiation growth factors, and replaced them for potential growth-inductive cytokines. However, flow cytometry measurements after 24 hours of induction did not result in significant increased levels of MESP1-mCherry expressing progenitors. In contrast, previous findings showed that NKX2-5-progenitor proliferation could be enhanced by IGF-1 or in combination with SAG (a smoothened agonist), and SB431542 (TGF-β/Activin/Nodal inhibitor)[9, 19]. The lack of response in our assay may indicate that a combination of different factors is required for MESP1 progenitors to proliferate, and possibly in the presence of the right microenvironment, such as the specific composition of ECM proteins or other cell types[26].

Proliferation of cardiomyocytes in the human heart has been described, although the rate is very limited[27]. It is important to understand how, if any, factors may induce cardiomyocyte proliferation, which may have an impact on treatment of heart failure by regeneration of the damaged heart after infarction[12-15, 28, 29]. Further optimization and technical advancement of screening assays, such as the combination of an Anillin reporter with a cardiac reporter appropriate for live imaging and automated quantification may prove to be very useful for identifying molecules and their underlying mechanisms for specific differentiation and differentiation of cardiac cell- (sub)types, which may lead to new therapeutic strategies for the treatment of heart failure.

FIGURES

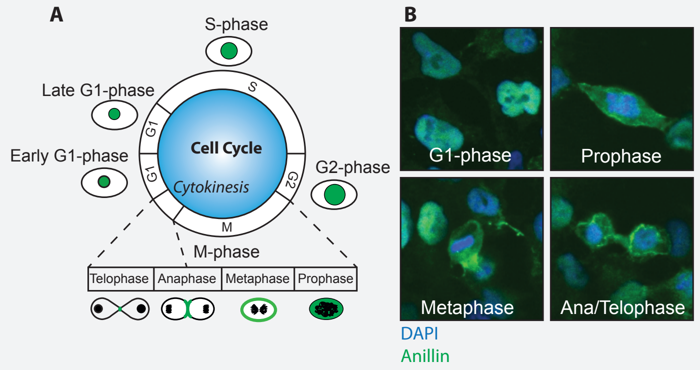


Figure 1. A: A schematic representation of Anillin. Localization of Anillin is visualized by fusion to eGFP and can be observed in the nucleus during G1-phase and migrates into the cytoplasm during prophase. During metaphase it appears in the cytoplasmic ring. Eventually, a midbody is formed during ana/telephase. Images are acquired using a 40x oil imm. Objective. *Figure 1a. is adapted from Hesse et al. Nature Communications 2011.* **B:** eGFP-Anillin localization in proliferating hESCs.

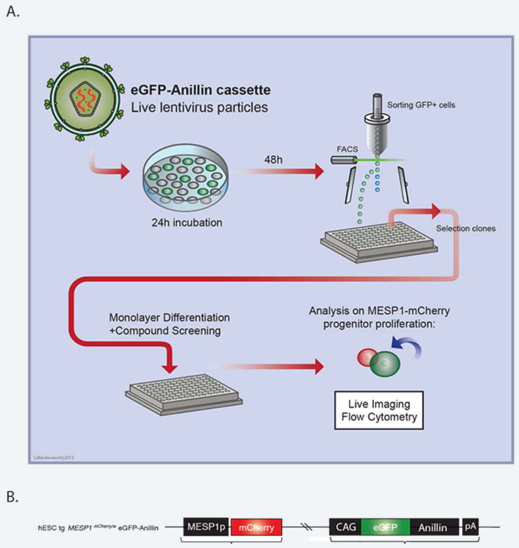


Figure 2. A: Overview scheme of how the eGFP-Anillin-MESP1mcherry/w reporter line was generated. hESCs were transfected with lentiviral particles containing the CAG-eGFP-Anillin cassette and were incubated for 24 hours. 48 hours after incubation, cells were sorted on eGFP expression and positive clones were selected for cardiac monolayer differentiations and compound screenings. Proliferation of MESP1-mCherry progenitors was assessed through live imaging of the eGFP-Anillin localizations**. B:** Schematic representation of the dual cardiac reporter line.

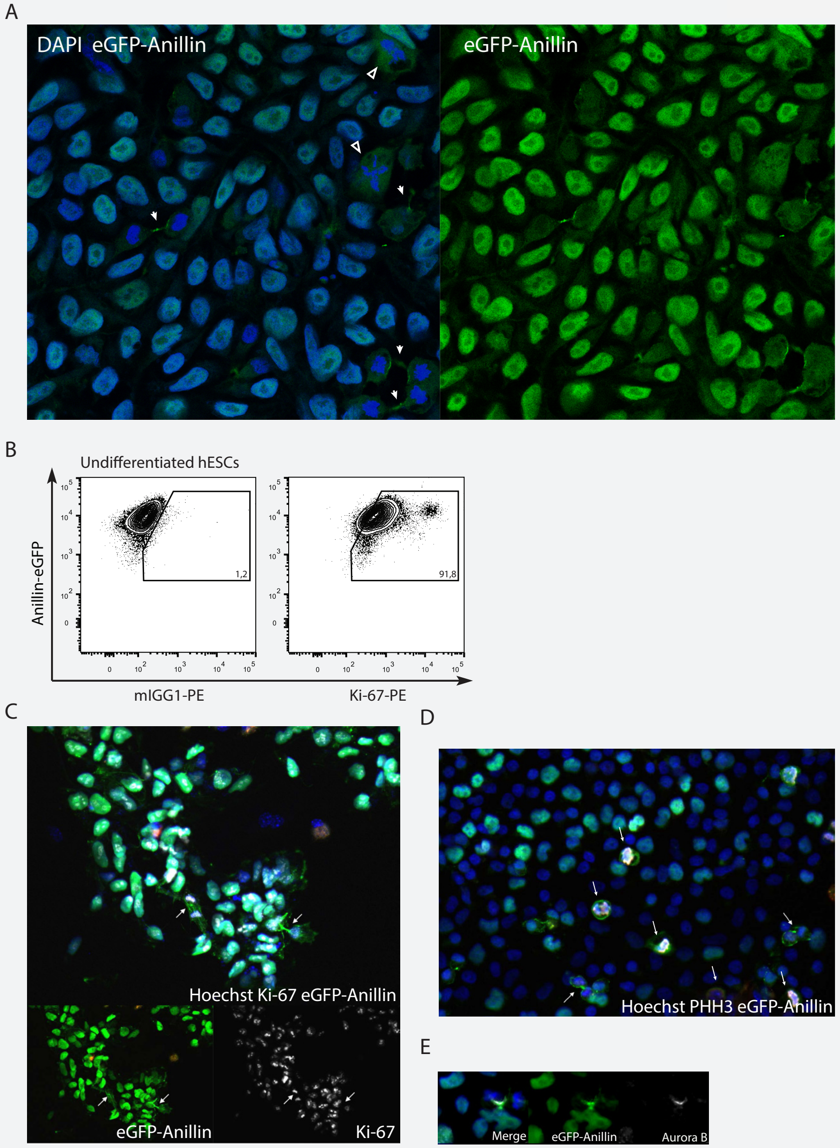


Figure 3. A: Picture of the dual reporter eGFP-Anillin-MESP1mcherry/w hESC line expressing eGFP–anillin under control of the CAG promoter. eGFP–Anillin (green) localization in the cytoplasm (open arrowhead), and midbody (arrow). **B:** Flow cytometric analysis of eGFP-Anilin hESCs stained for Ki-67 (anti-Ki-67-PE) and control isotype (mIGG1-PE). **C,D,E:** Stainings for the proliferation markers PHH3 (red), Aurora B (grey), and Ki-67 (grey) reveals co-expression with the eGFP-Anillin signal (green). Nuclei are stained with Hoechst dye (blue).

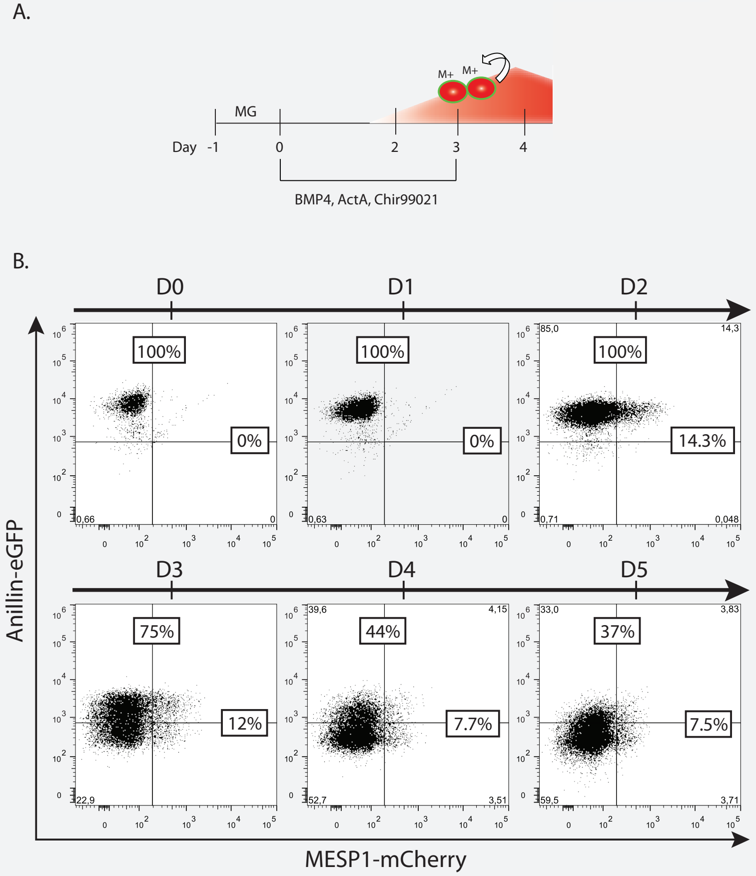


Figure 4. A: Cardiac monolayer differentiation protocol. MESP1-mCherry expression starts around day 2, and peaks between day 3 and day 4. Proliferation can be studied based on eGFP-Anillin expression. **B:** Flow cytometric analysis during cardiac differentiation, every 24 hours, from day 0 until day 5. eGFP-Anillin levels slowly decrease upon further differentiation.

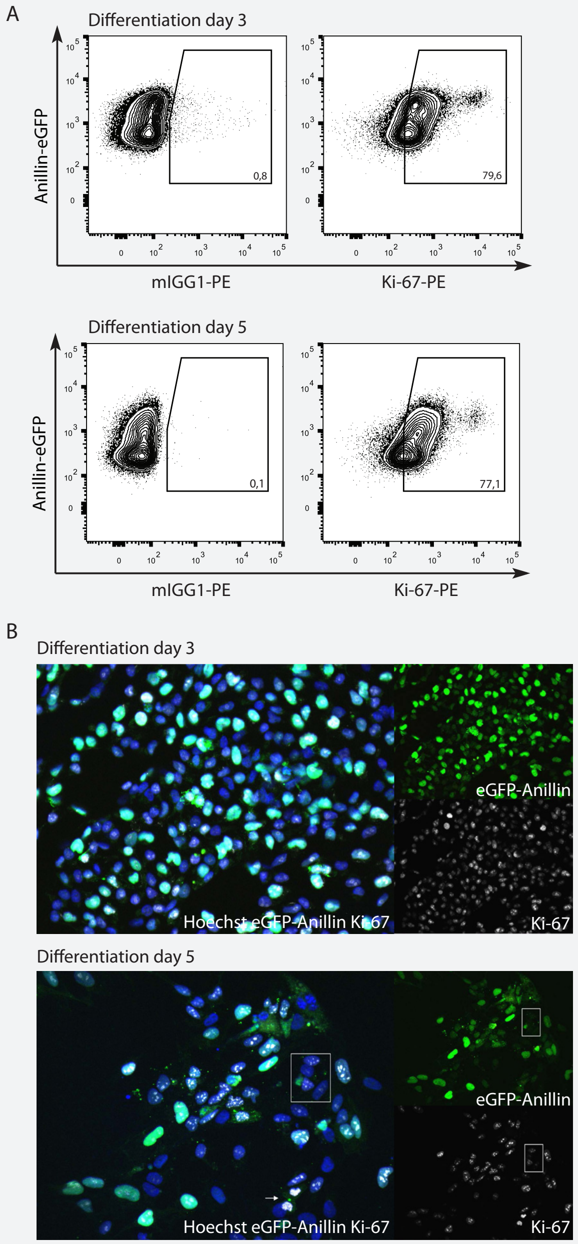


Figure 5. Flow cytometric analysis of Ki-67 co-expression with eGFP-Anillin- MESP1-mCherry expressing progenitors, at day 3 and day 5 of differentiation. Ki-67 levels remain high, despite a slight decrease, whereas eGFP-Anillin levels have decreased significantly, when compared to day 0 levels (Figure 4). **B:** Co-stainings for Ki-67 at day 3 and day 5 of differentiation revealed a large percentage of Ki-67 positive cells (active cell cycle) that show diminished eGFP-Anillin levels, indicating silencing of the construct.

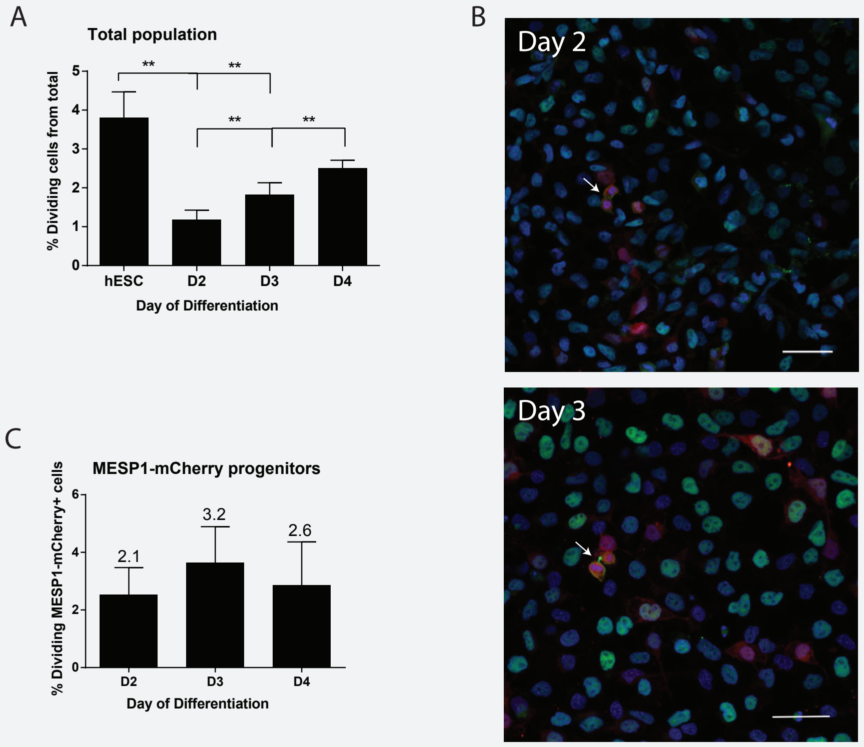
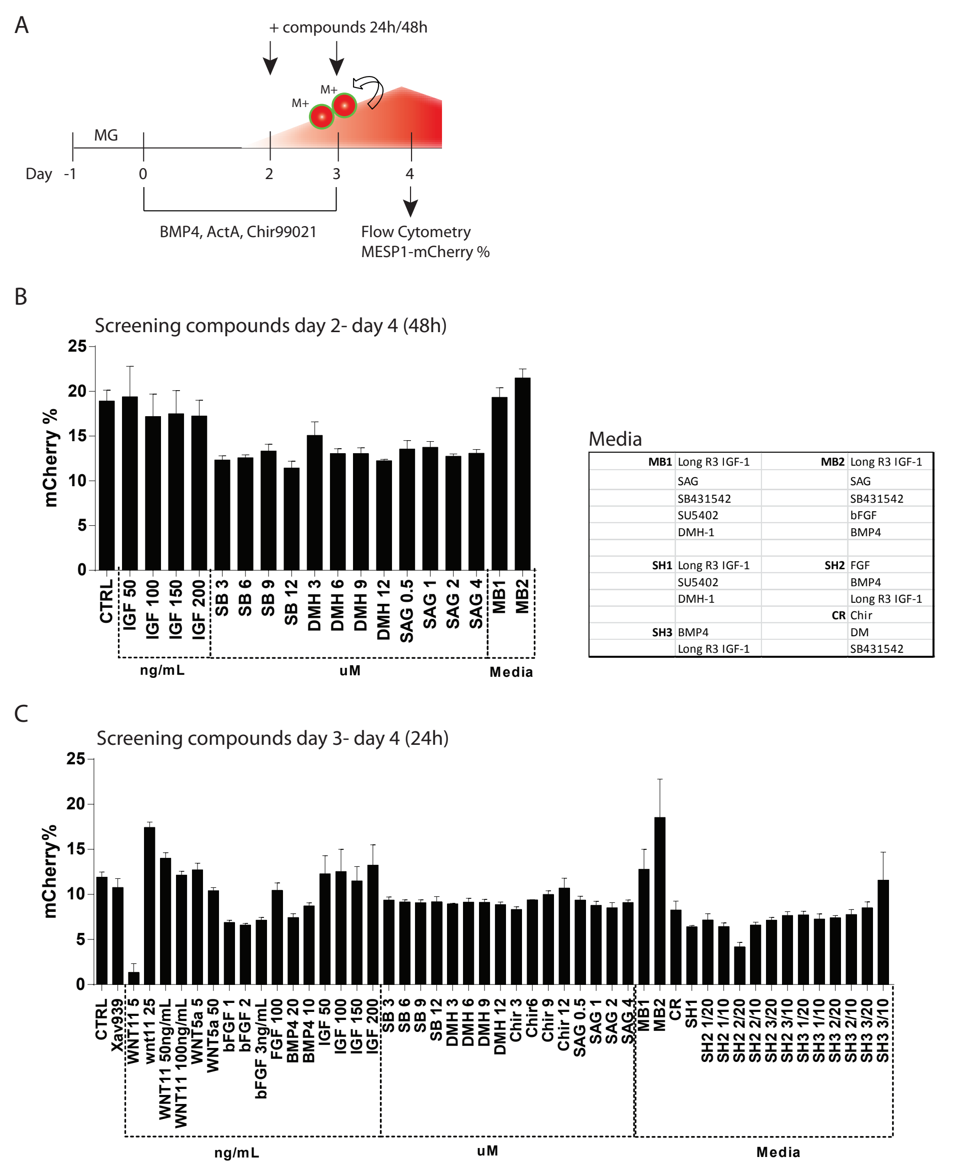


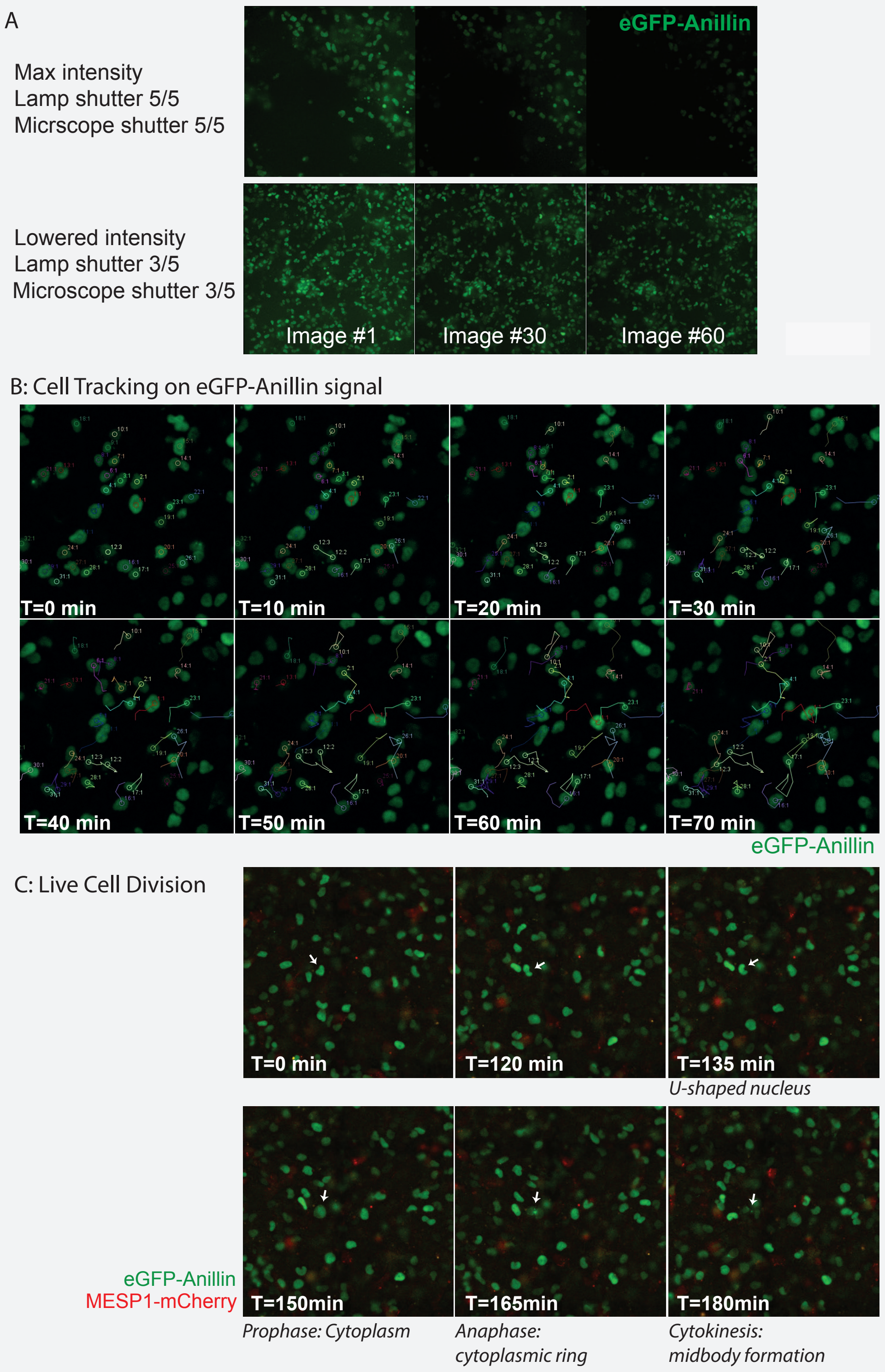
Figure 6. A: Percentage of all cells dividing (eGFP-Anillin M-phase localizations) upon early cardiac differentiation. **B:** Picture of a cardiac monolayer differentiation at day 3 visualizing proliferating (green) MESP1-mCherry progenitors (red). Nuclei are stained with DAPI (blue). **C:** Percentages of MESP1-mCherry progenitors dividing at day 2, day 3, and day 4 of differentiation. No significant differences between early and late progenitors were found. Scale day 2: 40uM. Scale day 3: 30uM.



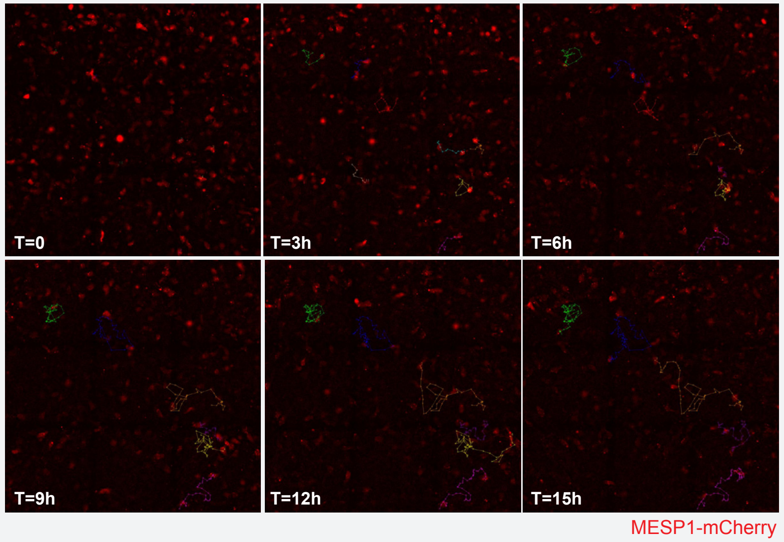
Media

MB1	Long R3 IGF-1	**MB2**	Long R3 IGF-1
	SAG		SAG
	SB431542		SB431542
	SU5402		bFGF
	DMH-1		BMP4
SH1	Long R3 IGF-1	**SH2**	FGF
	SU5402		BMP4
	DMH-1		Long R3 IGF-1
		CR	Chir
SH3	BMP4		DM
	Long R3 IGF-1		SB431542

Figure 7. A: Schematic overview of proliferation compound screenings upon a monolayer cardiac differentiation. Cells were differentiated as described before. At day 2 or day 3, BMP4, Activin A, and Chir were removed, and replaced with screening compounds for 48 or 24 hours, respectively. At day 4 of differentiation, the percentage of MESP1-mCherry progenitors in the total population was measured by flow cytometry. **B,C:** Percentages of MESP1-mCherry progenitors measured by flow cytometry at day 4 of differentiation, after treatment with compounds for 24 or 48 hours. No significant differences could be identified, when compared with control treatment (no GFs).



D: Cell Tracking of MESP1-mCherry expressing progenitors



E: Live Cell Division of MESP1-mCherry expressing progenitors

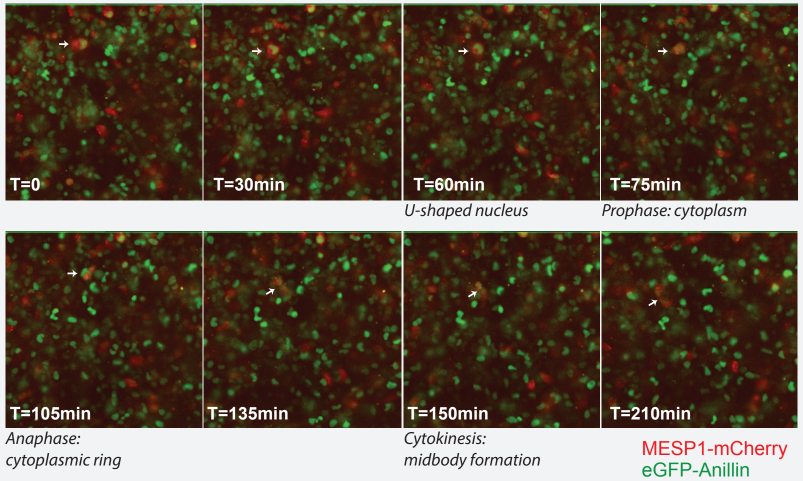


Figure 8: A: Settings for live-imaging fluorescent microscopy were optimized to prevent bleaching. 60 Sequential images were acquired with 63x glycerine immersion objective at minimised interval. The intensity of the lamp, exciting the fluorophores, is of great importance to prevent bleaching. By closing the lamp and microscope light shutter, the intensity of the light was reduced which resulted in decreased bleaching. **B:** A time interval of 10 minutes allows manual cell tracking of eGFP-Anillin expressing cells. **C.** Live visualization of dividing cell upon cardiac differentiation. A U-shaped nucleus is visible prior to cell division (arrow). Images are 2,5xzoom from original acquired image size. **C:** Manual tracking of MESP1-mCherry progenitors shows minimal migrative behaviour over a time span of 15 hours. Cell tracking stopped when cells were out of focus. Time interval of image acquisition: 15 minutes. **D:** Live visualization of dividing MESP1-mCherry progenitor (arrow). A U-shaped nucleus is visible prior to cell division. Images are 2,5xzoom from original acquired image size.

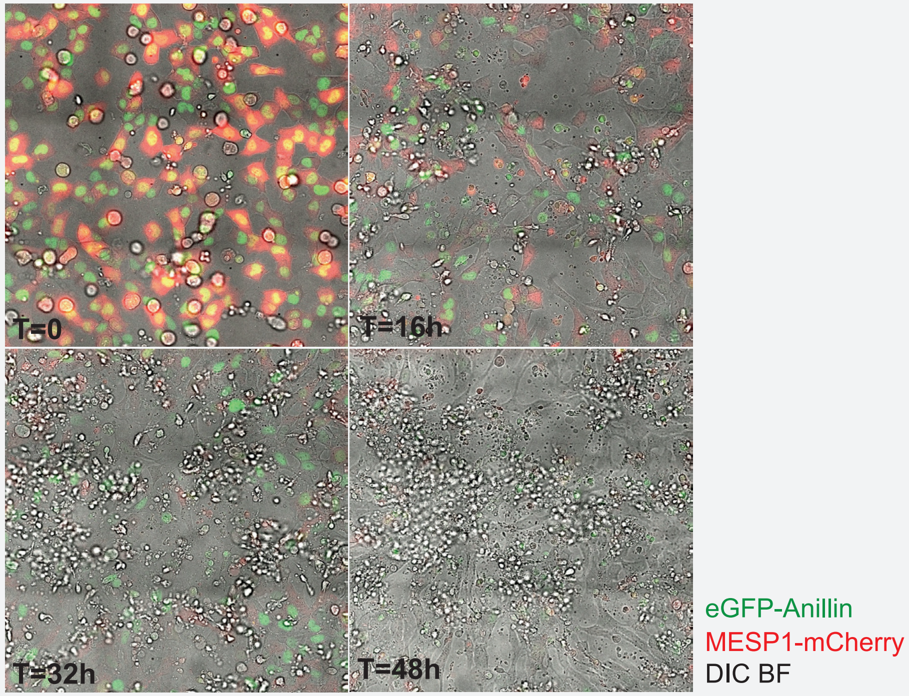


Figure 9. A: Time-lapse imaging of MESP1-mCherry sorted and replated progenitors. Images were acquired using a glycerol immersion 63x objective. T is given in minutes. To be able to track the highly migrative cells, a 10 minute interval was set. **B:** Time-lapse imaging of MESP1-mCherry sorted and replated progenitors. The Anillin-eGFP signal gets silenced upon further differentiation. mCherry levels decrease, as MESP1 levels are downregulated upon further differentiation. Morphological changes can be observed: cells become elongated and obtain a mesenchymal phenotype. T is given in hours.

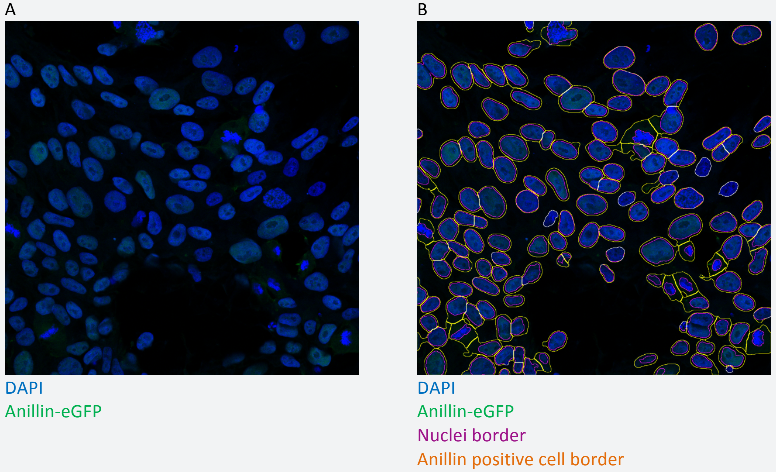
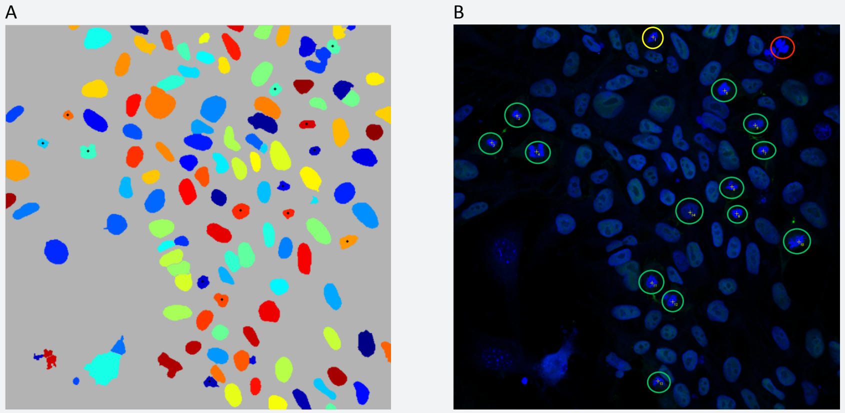
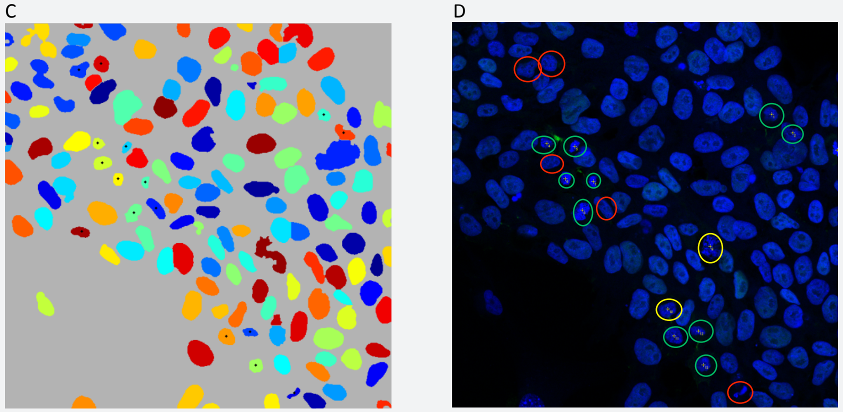


Figure 10. Automated Segmentation Result.



Automated annotation 1

Manual annotation 1



Automated annotation 2

Manual annotation 2

E. Comparison of automated and manual cell segmentation by F-score analysis.

	TP:	TP + FP:	TP + FN:
Total:	399	490	444
	Precision:	Recall:	F-score:
	0,814285714	0,898648649	0,854389722

$$\text{Precision} = \frac{\text{TP}}{\text{TP + FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP + FN}}$$

$$\text{F-score} = \frac{\text{2PR}}{\text{P + R}}$$

Figure 11. Automated Annotation versus Manual Annotation. To identify the degree of errors in the automated analysis, we compared the results from manually quantifications of 22 images with their automated quantified counterpart. **A-D:** Green circles: positive and selected. Yellow circles: false negative. Red circles: false positive. **E:** Calculation of F-score. TP: true positives. FP: false positives. FN: false negatives. P: Precision. R: Recall.

Weighted classification error: the sum of k-th class error divided by the k-th class size multiplying the weight k which is the prior probability of k-th class.

$$error_w = \sum_{k=1}^{i} w_k * (\frac{e_k}{n_k})$$

e_k is the number of misclassified training examples from the k-th class and n_k is the size of the k-th class in the training set.

Calculation of weighted classification error in this study:

$$error_w = \left(\frac{291}{1500}\right) * \frac{e_{div}}{n_{div}} + \left(\frac{1209}{1500}\right) * \frac{e_{non-div}}{n_{non-div}}$$

Minimal error$_w$ calculated for each feature selection method:

	Individual		B&B		Backward		Forward	
	min	sem	min	sem	min	sem	min	sem
LDC	0.0322	0.0014	0.0314	0.0013	0.0344	0.0013	0.0352	0.0012
QDC	0.0388	0.0017	0.0336	0.0014	0.0406	0.0014	0.0426	0.0013
KNNC	0.0359	0.0014	0.0256	0.0011	0.0317	0.0013	0.0320	0.0013
SVC	0.0355	0.0013	0.0285	0.0013	0.0314	0.0014	0.0313	0.0013

Individual: best individual N feature selection method; B&B: branch and bound feature selection method; Backward: greedy backward elimination method; Forward: greedy forward selection method; min: minimal value of mean weighted error; sem: standard error of mean; LDC: the linear classifier; QDC: the quadratic classifier; KNNC: the k-nearest neighbour classifier; SVC: support vector machine classifier. REFS: www.mathworks.com and Polo 1997.

Table 1. Minimal Value of Mean Weighted Errors for each Feature Selection Method.

Cytokine	Concentration
Long R3 IGF-1	50 ng/mL
	100 ng/mL
	150 ng/mL
	200 ng/mL
SB431542 (Tocris)	3 uM
	6 uM
	9 uM
	12 uM
Dorsomorphin	3 uM
Homologue-1	6 uM
(DMH)	9 uM
	12 uM
SAG (Millipore)	0.5 uM
	1 uM
	2 uM
	4 uM
Wnt11 (R&D)	5 ng/mL
	25 ng/mL
	50 ng/mL
	100 ng/mL
bFGF	1 ng/mL
	2 ng/mL
	3 ng/mL
	100 ng/mL
Wnt5a (Millipore)	5 ng/mL
	50 ng/mL
XAV939	5 uM

Medium	FGF (ng/mL)	BMP4 (ng/mL)	IGF
SH2	1	10	100 ng/mL
	2	10	100 ng/mL
	1	20	100 ng/mL
	2	20	100 ng/mL
	3	10	100 ng/mL
	3	10	100 ng/mL
SH3	1	10	-
	2	10	-
	1	20	-
	2	20	-
	3	10	-
	3	10	-

Medium	IGF	SAG (Millipore)	SB431542 (Tocris)	SU5402	DMH1	bFGF	BMP4
MB1	100 ng/mL	1 uM	5 uM	5 uM	0.5uM	-	-
MB2	100 ng/mL	1 uM	5 uM	-	-	1 ng/mL	20 ng/mL

Medium	Chir	DM	SB431542
CR	3 uM	2uM	0.5uM

Supplemental Table 1. Microscopy live imaging settings

Objective:	63x Glyc
Channels;	eGFP, mCherry and Brightfield
Image Interval:	10 minutes
Imaging Time:	64 hours
Z-Stack:	3x 5μM
Binning:	2x2
Autofocus:	Brightfield, compensated with zstack
Stitching:	3x3 (x: 59.6 y: 35.371)

Channels:	Exposure (ms)	EM gain	Intensity
eGFP	125	1100	4,3
mCherry (txred)	125	1100	4,3
Brightfield	30	1100	5

Supplemental Table 2. Small molecules/Growth factors screened

REFERENCES

1. Fuster V, Voute J, Hunn M, et al. Low priority of cardiovascular and chronic diseases on the global health agenda: a cause for concern. CIRCULATION … 2007;116(17):1966–1970.
2. Orlova VV, van den Hil FE, Petrus-Reurer S, et al. Generation, expansion and functional analysis of endothelial cells and pericytes derived from human pluripotent stem cells. NATURE PROTOCOLS 2014;9(6):1514–1531.
3. Burridge PW, Matsa E, Shukla P, et al. Chemically defined generation of human cardiomyocytes. NAT METH 2014:–.
4. Hartogh Den SC, Schreurs C, Monshouwer-Kloots JJ, et al. Dual Reporter MESP1 mCherry/w-NKX2-5 eGFP/whESCs Enable Studying Early Human Cardiac Differentiation. STEM CELLS 2014;33(1):56–67.
5. Elliott DA, Braam SR, Koutsis K, et al. NKX2-5eGFP/w hESCs for isolation of human cardiac progenitors and cardiomyocytes. NAT METH 2011;8(12):1037–1040.
6. Qyang Y, Martin-Puig S, Chiravuri M, et al. The Renewal and Differentiation of Isl1+ Cardiovascular Progenitors Are Controlled by a Wnt/β-Catenin Pathway. CELL STEM CELL 2007;1(2):165–179.
7. Engels MC, Rajarajan K, Feistritzer R, et al. IGF promotes cardiac lineage induction by selective expansion of cardiogenic mesoderm in vitro. EUROPEAN HEART JOURNAL 2013;34(suppl 1):P5037.
8. Kwon C, Qian L, Cheng P, et al. A regulatory pathway involving Notch1/β-catenin/Isl1 determines cardiac progenitor cell fate. NATURE CELL BIOLOGY 2009;11(8):951–957.
9. Birket MJ, Ribeiro MC, Verkerk AO, et al. Expansion and patterning of cardiovascular progenitors derived from human pluripotent stem cells. NATURE BIOTECHNOLOGY 2015;33(9):970–979.
10. Saga Y. MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube. DEVELOPMENT 1999:1–11.
11. Hesse M, Raulf A, Pilz G-A, et al. Direct visualization of cell division using high-resolution imaging of M-phase of the cell cycle. NAT COMMS 2012;3:1076.
12. Engel FB, Hsieh PCH, Lee RT, et al. FGF1/p38 MAP kinase inhibitor therapy induces cardiomyocyte mitosis, reduces scarring, and rescues function after myocardial infarction. PROC. NATL. ACAD. SCI. U.S.A. 2006;103(42):15546–15551.
13. Kühn B, del Monte F, Hajjar RJ, et al. Periostin induces proliferation of differentiated cardiomyocytes and promotes cardiac repair. NAT. MED. 2007;13(8):962–969.
14. Xin M, Olson EN, Bassel-Duby R. Mending broken hearts: cardiac development as a basis for adult heart regeneration and repair. NAT REV MOL CELL BIOL 2013;14(8):529–541.
15. Bersell K, Arab S, Haring B, et al. Neuregulin1/ErbB4 Signaling Induces Cardiomyocyte Proliferation and Repair of Heart Injury. CELL 2009;138(2):257–270.
16. Bass GT, Ryall KA, Katikapalli A, et al. Automated image analysis identifies signaling pathways regulating distinct signatures of cardiac myocyte hypertrophy. JOURNAL OF MOLECULAR AND … 2012.
17. Webb AR. Statistical Pattern Recognition - Andrew R. Webb - Google Books. 2003.
18. Ruusuvuori P, Äijö T, Chowdhury S. Evaluation of methods for detection of fluorescence labeled subcellular objects in microscope images. BMC … 2010.
19. Engels MC, Rajarajan K, Feistritzer R, et al. Insulin-Like Growth Factor Promotes Cardiac Lineage Induction In Vitro by Selective Expansion of Early Mesoderm. STEM CELLS 2014;32(6):1493–1502.
20. Juan J, Cubero C. Weighted Classification Using Decision Trees for Binary Classification Problems. LECTURE NOTES IN COMPUTER SCIENCE 1997.
21. Abdelsamea M, Mohamed MH, Bamatraf M. An Effective Image Feature Classiffication using an improved SOM. ARXIV 2015.
22. García-Olalla O, Alegre E, Fernández-Robles L, et al. Vitality Assessment of Boar Sperm Using an Adaptive LBP Based on Oriented Deviation. In: Computer Vision-ACCV ….Vol 7728. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013:61–72.
23. Zavitz KH, Zipursky SL. Controlling cell proliferation in differentiating tissues: genetic analysis of negative regulators of G1-->S-phase progression. CURR. OPIN. CELL BIOL. 1997;9(6):773–781.
24. Xia X, Zhang Y, Zieth CR, et al. Transgenes Delivered by Lentiviral Vector are Suppressed in Human Embryonic Stem Cells in A Promoter-Dependent Manner. STEM CELLS AND DEVELOPMENT 2007;16(1):167–176.
25. Pre Du BC, Van Veen TAB, Young ME, et al. Circadian Rhythms in Cell Maturation. PHYSIOLOGY 2014;29(1):72–83.
26. Ieda M, Tsuchihashi T, Ivey KN, et al. Cardiac fibroblasts regulate myocardial proliferation through beta1 integrin signaling. DEVELOPMENTAL CELL 2009;16(2):233–244.
27. Bergmann O, Zdunek S, Felker A, et al. Dynamics of Cell Generation and Turnover in the Human Heart. CELL 2015;161(7):1566–1575.
28. Hartogh SCD, Sluijter JP, Doevendans PA, et al. Cellular Therapy for the Infarcted Myocardium. In: Translational …. Totowa, NJ: Humana Press; 2012:341–390.
29. Passier R, van Laake LW, Mummery CL. Stem-cell-based therapy and lessons from the heart. NATURE 2008;453(7193):322–329.

CHAPTER 8

General discussion

It is critical to gain knowledge in the underlying mechanisms that control human cardiovascular developm ent, which helps us to understand the onset of congenital cardiovascular diseases, and to develop optimal culture methods for efficient *in vitro* cardiomyocyte differentiation from hPSCs, which are of interest for final translational applications including screening and efficacy assays for disease modelling, drug discovery and development, personalized medicine, and perhaps the regeneration of cardiovascular tissues for therapeutic purposes. In this thesis, we show how genetic manipulation of human pluripotent stem cells (hPSCs), resulting in the genomic integration of a fluorescent protein encoding sequence at the locus of a key cardiac transcription factor, allows us to visualize and isolate early pre-cardiac progenitors subpopulations, and to study the molecular mechanisms involved in their further differentiation to cells of the cardiac lineage, including smooth muscle cells, endothelial cells, and cardiomyocyte subtypes.

IN VITRO MODELLING OF HUMAN CARDIAC DEVELOPMENT AND DIFFERENTIATION USING HUMAN PLURIPOTENT STEM CELLS:

HOW DO FLUORESCENT GENE REPORTERS CONTRIBUTE?

Temporal gene expression analysis upon cardiac differentiation

Multiple fluorescent PSC reporter lines have been generated to visualize the derivation of cardiomyocytes and to isolate PSC-derived cardiac cells, including cardiac progenitor populations and subtypes of cardiomyocytes (reviewed in **Chapter 2**). Most *in vitro* fluorescent reporter PSC lines described in literature so far, contain a construct that follows the expression of a single gene during differentiation, such as cardiomyocyte reporters NKX2-5$^{eGFP/w}$, NCX-1-eGFP, or alpha-MHC-mCherry human embryonic stem cell (hESC) lines[1, 2] [3]. These lines allow visualization and isolation of the reporter-expressing cells, cardiac progenitors and cardiomyocytes. On the basis of a fluorescent marker, this allows screening molecules for their role in cardiomyocyte differentiation or the identification of subtype specific cell surface markers. This was used for example to identify SirpA and VCAM1 as cell surface markers for early cardiomyocytes[3, 4]. In **Chapter 3** we describe the generation of a dual fluorescent cardiac reporter MESP1^{mCherry}/wNKX2-5$^{eGFP/w}$ hESC line, that allows the derivation of early pre-cardiac mesoderm progenitors to be monitored and their further lineage-committed differentiation to the cardiac lineage, marked by NKX2-5-eGFP expression. Comprehensive gene expression analysis of MESP1 cardiac derivatives in **Chapter 4** showed how cardiac differentiation of these progenitors recapitulates the early steps of *in vivo* heart development, through the sequential expression of key cardiac transcription factors and the expression of functional and structural genes in definitive cardiomyocytes, providing evidence of how such a dual fluorescent cardiac reporter line is useful for studying early human cardiac differentiation. Several previous studies have shown global gene expression patterns upon cardiac differentiation from hPSCs[5, 6]. However, in our study, we could follow the formation of cardiac cells, deriving from an early pre-cardiac-committed lineage, and decreasing the risk for upregulation of genes that are not associated with the cardiac lineage, and thus increasing the predictive value of this model. Putative (co)-regulators of cardiac lineage commitment were identified based on their temporal enrichment levels compared to MESP1 negative derivatives, their DNA-binding properties, and their predicted interaction with other established key regulators. We identified homeobox domain containing HOXB2 as potential interactive transcriptional partner for MEIS1, MEIS2, and PBX3. This was strengthened by a similar expression level pattern throughout cardiac differentiation. MEIS1, MEIS2, and PBX3 all share regulating roles in either heart looping and chamber septation stages, and cardiac lineage commitment. Moreover, MEIS1, MEIS2, HOXB2, and PBX3 were all upregulated upon Mesp1 induction in mouse ESCs, indicating that they all act downstream of Mesp1 (REF). Other DNA binding transcription factors that showed significant stage-specific expression levels upon cardiac differentiation, included zinc finger protein containing transcription factors ZNF503, ZFHX3, ZFPM1, and ZFPM2, that may have co-regulatory roles for established cardiac transcription factors or repressive roles on transcription of key genes involved in other lineages, such as skeletal[7]. Such a repressive role of other lineages could perhaps be of importance for enhancing cardiac lineage commitment.

Additionally, in our study, shortly after temporal MESP1 expression, we also identified significant enriched temporal expression levels of DNA binding genes with an unknown role in heart development at present, including ZBTB16, RUNX1T1, and TSHZ2. The exact role of such putative regulators, which may include lineage commitment, but could also govern cell cycle regulation, proliferation, and/or survival, has to be determined in functional studies, such as *in vitro* inducible knock down studies in MESP1-lineage specific derivatives, or downstream target gene analysis using chip sequencing. Functional studies are a prerequisite for definite conclusions about gene function and role.

In the study of Paige et al.[5], they show how temporally regulated gene expression levels are accompanied by programmed temporal alterations in chromatin structure, which distinguishes key regulators of cardiovascular development from other genes. Currently, further studies are being performed to look into chromatin patterns and dynamics of genomic regions of transcription factors that we identified in this thesis, which would strengthen their role in cardiac differentiation and/or lineage commitment. This will be discussed briefly below.

Future Directions: integrative (epi)genetic analysis upon cardiac lineage commitment.

Despite increasing knowledge on key cardiac transcriptional regulators, signalling molecules and interacting pathways, the role of epigenetic alterations, such as chromatin dynamics and histone modification patterns, are largely unknown, but are proposed to play crucial roles in gene regulation during mammalian heart development and differentiation[5, 8].

Recent studies have indicated that large-scale reorganization of the genome, and epigenetic mark transitions, such as histone modification patterns, coincide with the commitment of pluripotent cells to differentiation[5, 8, 9]. A large integrative study on (epi)genetic and transcriptional regulation during early cardiac differentiation, dual fluorescent cardiac gene reporter MESP1$^{mCherry/w}$-NKX2-5$^{eGFP/w}$ hESCs would therefore be useful, and is currently being performed. It allows the isolation of defined cardiac (progenitor) populations in order to study gene- and cell type-specific chromatin dynamics, histone modifications patterns, RNA expression levels, and enhancer regions. The increasing development and technical improvements of molecular genomic technologies, such as ATAC-sequencing (studying chromatin availability), ChiP-sequencing, and Chromosome Conformation Capture Hi-C (studying chromatin conformations) will allow us to integrate large genomic and epigenetic information at stage-specific cardiac cell types, in order to contribute to the identification of stage-specific (distal) enhancer- and/or promoter elements and unidentified (co)-regulators of cardiac lineage commitment, and potentially also of cardiac subtype diversification. Current pitfalls that require improvement: high costs are made, and high cell numbers are required for such studies. Nonetheless, interest into large integrative (epi)genetic studies is growing.

Screening signaling molecules upon cardiac lineage commitment

From a large number of differentiation studies we know that the inhibition of Wnt-pathway signalling, using distinct Wnt-pathway inhibiting molecules such as Xav939, is crucial for efficient differentiation of pre-cardiac mesoderm towards the cardiac lineage[10-12]. MESP1 expressing mesoderm progenitors may differentiate towards distinct lineages, including skeletal, hematopoietic, and cardiac lineages[13]. In order to determine how the presence and/or absence of distinct signaling molecules may induce pre-cardiac MESP1 expressing mesoderm stage progenitors to further segregate towards the cardiac lineage, dual fluorescent cardiac gene reporter MESP1$^{mCherry/w}$-NKX2-5$^{eGFP/w}$ hESCs is of excellent utility. In **Chapter 3**, we studied a selection of signaling molecules, which included small molecule Wnt-pathway inhibitor Xav939[12], TGF-β/act/Nodal antagonist SB431542, and BMP inhibitor Dorsomorphin. Besides the strong role for Wnt-pathway inhibition in efficient cardiomyocyte differentiation from mesoderm stage, none of the other inhibitors led to an increase in NKX2-5-eGFP expression, where TGF-β/act/Nodal inhibition even seemed to decrease cardiomyocyte differentiation, in contrast with previous literature describing an inhibitory role for TGF-β/act/Nodal signaling in mesoderm to cardiomyocyte differentiation[14, 15]. Such diverse effects could be highly dependent on ligand or small molecule concentration and timing, and in the case of growth factors, specific activity, which is rarely measured. In addition, growth factor effects on cardiomyocyte sublineage commitment, may depend on the combinatorial presence of other cytokines so that it will be necessary to study the effects of a large number of growth factors or small molecules at middle- or high-throughput level (in 96-, 384- or 1536 well plates). For *in vitro* differentiation assays, the additional use of cardiomyocyte subtype reporters, resulting in a multi-color differentiation study may be ideal. This would allow us to follow hPSC differentiation to mesoderm, followed by cardiac specification, and definite subtype derivation. Examples of cardiomyocyte subtype reporters, established in the NKX2-5$^{eGFP/w}$ hESC line, are discussed below (**Chapter 6 and 7**).

The search for expandable early cardiac progenitor populations for translational applications

In order to generate large numbers of cardiomyocyte subtypes for many translational purposes, it would be advantageous to have cardiac lineage-committed progenitor populations, which can be expanded and further differentiated, under tight control, to large pure numbers of the required cardiac subtype. However, only a limited number of studies has shown the ability to maintain cardiac progenitors in their self-renewing developmental state, most with the limitation of a single expansion (reviewed in Birket et al.)[16].

There is a strong demand for the identification of specific cell surface markers that identify and can be used for isolation of hPSC-derived cardiac progenitor populations. In this thesis, we have identified the presence of previously described cardiac progenitors, based on ROR2, CD13 (ANPEP), PDGFRA and KDR, and SirpA and VCAM1 expression[4, 17]. However, the search for additional early cardiac-specific surface markers is still required, since ROR2, CD13, PDGFRA, and KDR showed a more broad mesoderm expression (**Chapter 3**). From our micro array analysis of MESP1-lineage derivatives in **Chapter 4**, we suggest potential novel surface marker proteins, based on their temporal-specific upregulated transcript levels. It will now be of importance to study their expression dynamics at protein level by antibody-binding experiments. Based on current knowledge of cardiac progenitor biology, a combinatorial expression of multiple surface markers may identify cardiac-lineage specific progenitors[4, 18, 19]. Thus, the exact genes and surface markers identifying critical cardiac progenitor populations remain to be defined. The next step will be to study their expandable capacity. In **Chapter 7,** we attempted to study the proliferative capacity of MESP1-mCherry expressing progenitors. A previous study described how they could expand an early hPSC-derived cardiovascular progenitor population (expressing MESP1). Their expansion protocol was based on the inhibition of BMP and activin/nodal signalling, and the activation of Wnt signalling[20, 21]. Even though both progenitors seem to be at a similar developmental stage, we could not reproduce MESP1 progenitor maintenance in our hands. Moreover, we attempted a variety of other cytokines, including those that were described to be involved in NKX2-5 progenitor maintenance[21].

The expansion capacity of cardiac progenitors may be based on their developmental stage and the extracellular niche that they may require. More studies into signals required for cardiac progenitor expansion will be necessary before they can actually be ready for translational purposes.

FLUORESCENT GENE REPORTERS FOR CARDIOMYOCYTE SUBTYPE DERIVATION

The identification of molecular regulators involved in cardiomyocyte subtype derivation.

In **chapter 5 and 6**, we describe the generation of two separate dual cardiomyocyte fluorescent hPSC reporter lines, either for the identification of ventricular or nodal-like cardiomyocytes, and to study molecular regulators important for their derivation. Both lines were generated in the previously described NKX2-5$^{eGFP/w}$ hESC line[3]. In chapter 6, we generated a reporter construct in which an atrioventricular-specific enhancer and a pan-cardiac enhancer, that were previously identified to synergistically activate TBX3 expression in the atrioventricular conduction system of the developing mouse embryo, were coupled to a promoter element and fluorescent protein mCherry. This reporter construct was randomly integrated into the genome of NKX2-5$^{eGFP/w}$ hESCs through lentiviral delivery.

The use of gene regulatory elements that have been identified through *in vivo* studies in other organisms, such as rodents, the zebrafish, and chick, is a common strategy for *in vitro* fluorescent reporter studies[22-24]. Likewise used the group of Laflamme, in order to identify PSC-derived nodal-like cardiomyocytes, a proximal promoter-enhancer region of the chicken GATA6 (cGATA6) gene that was earlier identified to be selectively activated in the atrioventricular (AV) node and the bundle of His of the adult mouse heart[24, 25]. hESC-cardiomyoycte cultures were transduced with a lentiviral vector in which cGATA6 drives expression of enhanced green fluorescent protein (EGFP), resulting in the identification of cardiomyocytes with a nodal-like phenotype, evidenced by their action potential. Furthermore, they showed how ErbB antagonist AG1478 could enhance the differentiation of hESC towards nodal-like cardiomyocytes. Therefore, in our study, we exposed dual reporter TBX3e-mCherry-NKX2-5$^{eGFP/w}$ hESCs to AG1478, shortly after mesoderm formation. However, we failed to identify TBX3e-mCherry expressing cells. Also preliminary screening with a variety of other cytokines and small molecules did not show any positive cells. Although we think that the combinatorial use of both TBX3 enhancers with cardiomyocyte marker NKX2-5 is of interest to study molecules involved in the derivation of nodal-like cardiomyocytes, we realized that our strategy could be biased by many biological and technical issues, including transgene silencing, mouse/human differences, or lack of activity of both enhancers in our system (technical issues are described below in more detail). It would be of interest to study a variety of other CCS-specific regulatory elements, or to report endogenous TBX3 expression through a knockin or fusion strategy.

In **chapter 6**, we aimed to monitor and isolate hESC-derived ventricular cardiomyocytes by the generation of dual cardiomyocyte reporter MYL2$^{w/T2A\text{-}mCherry}$-NKX2-5$^{eGFP/w}$ hESCs through an endogenous fusion strategy. Although from electrophysiology data on cardiomyocyte cultures we know that PSC differentiation often results in a predominantly ventricular fate, we could identify MYL2-mCherry expressing cardiomyocytes, only at low levels. This agrees with previous studies using mouse and human MYL2 PSC reporter lines[22, 26] and from our microarray analysis of MESP1 derivatives (**Chapter 3**) we know that MYL2 starts to be significantly expressed around day 14 of differentiation. Moreover, as MYL2 levels in the adult heart are >2 fold upregulated compared with hPSC-CMs, and as MYL2 levels in hPSC-CMs increase upon maturation (Ribeiro et al.[27] and **Chapter 6**), MYL2 may also be an interesting marker for measuring the degree of ventricular cardiomyocyte maturation.

Although recent protocols, based on retinoic acid signalling, have been shown to drive hPSCs towards atrial fate[28, 29], fluorescent atrial reporter lines have not yet been developed. Genes of interest to target could include atrial-specific sarcolipin or COUPTF genes[29-31]. This would allow monitoring of a variety of other molecules possibly involved in atrial fate differentiation, and may contribute to optimize growth factor concentration and timing. The availability of high numbers of hPSC-derived atrial cardiomyocytes will be of interest for assay development for atrial-specific associated disease modelling and drug discovery and/or toxicity[29].

Another strategy to obtain atrial- or ventricular specific cardiomyocytes could be based on a panel of surface markers being expressed, without the requirement of genetic modifications of hPSC lines. The primary use of fluorescent reporter lines allows the isolation of cardiomyocyte specific cell types, which then may contribute to the identification of cell type specific surface markers that could be used for future subtype isolation.

The optimal strategy for the generation of a cardiac fluorescent reporter line should be carefully considered: biological/technical hurdles and new perspectives.

As mentioned before, prior to the generation of a fluorescent reporter line in hPSCs, it is important to consider the technical and biological hurdles that may be faced upon establishment of such a line. Comprehensive details are reviewed in **Chapter 2**. In this thesis, we show the generation of four distinct reporter lines, using different modification strategies. In **Chapter 3**, we generated a MESP1mCherry/w reporter line through the replacement of one allele with the coding sequence of fluorescent protein mCherry, allowing reporting the activity of the endogenous MESP1 promoter. mCherry is monomeric, has a reasonable brightness, and a fast maturing folding efficiency at 37 °C and a high photostability, optimal for long-term imaging studies[32]. Targeting was performed in the previously by Elliott et al. generated NKX2-5eGFP/w hESC line[3]. MESP1-mCherry expressing progenitors could be distinguished from NKX2-5-eGFP cardiac progenitors, due to their non-overlapping timing of expression, and the non-overlapping excitation and emission spectra of mCherry (excitation at 587 nm) and eGFP (excitation at 488 nm). In a similar matter, we made use of these distinguishing spectra when we generated the dual cardiac reporter hESC lines TBX3e-mCherry-NKX2-5$^{eGFP/w}$ in **Chapter 5** and MYL2$^{w/T2A\text{-}mCherry}$-NKX2-5$^{eGFP/w}$ in **Chapter 6.**

Although we could not detect molecular abnormalities upon cardiomyocyte differentiation from MESP1$^{mCherry/w}$NKX2-5$^{eGFP/w}$ hESCs, generating a knockin reporter line through gene replacement brings along the risk for haploinsufficiency defects. Therefore, we generated the MYL2w/T2A-$^{mCherry-}$NKX2-5$^{eGFP/w}$ hESC line through a fusion knockin strategy, where we made use of the small viral T2A fusion peptide[33], which, upon activation of the MYL2 promoter, results in the generation of one transcript, containing both MYL2 and mCherry coding sequences, and is later translated into two separate proteins. This latter strategy seems to have a great potential in stem cell fate reporter technology, especially with the eye on the generation of multicolour reporter lines. However, in hPSCs, there are few reports on the use of these small fusion peptides, thus technical hurdles may still exist. Although in our study, we show the presence of MYL2-mCherry expressing cardiomyocytes, the translation efficiency of both proteins was undefined. Another strategy would be to fuse a fluorescent protein directly the targeted gene, although this may affect gene function.

A less labour-intensive method to develop fluorescent reporter lines is the lentiviral delivery of a reporting construct. In **Chapter 5** lentiviral transfection of NKX2-5$^{eGFP/w}$ hESCs, resulted in the random integration of the TBX3e-Hsp68-mCherry reporting construct into the hESC genome. However, as mentioned before, lentiviral integration into active genome sequences is random, and can be highly susceptible to transgene silencing due to surrounding regulatory elements or upon cardiac lineage differentiation (shown in **Chapter 7**). Positive controls to select correctly targeted clones is a prerequisite, but also labour-intensive.

Therefore, stem cell reporter technology requires technical improvements, including more efficient and less labour-intensive locus-specific targeting strategies (perhaps through the use of the new advances in CRISPR/CAS technology), and through the identification of loci in the human genome that are not silenced upon cardiac-specific lineage differentiation[34-36]. However, it may not be a guarantee that such a "docking locus" may be active in every cardiac sub lineage that will be studied.

HUMAN PLURIPOTENT STEM CELL-DERIVED CARDIAC PROGENITORS AND CARDIOMYOCYTES: TRANSLATIONAL APPLICATIONS

The establishment of hPSC technology and advanced cardiomyocyte differentiation protocols has opened a new platform for cardiac development, tissue engineering, disease modelling, and drug efficacy and toxicity testing. There are still many hurdles that need to be taken before stem cell based therapies find its way in the clinic for the treatment of heart failure. In recent years the recognition for using hPSC-derived cardiomyocytes for disease modelling, drug discovery, and drug toxicity screenings has increased significantly. hPSC-technology aims to develop highly reliable animal-free pre-clinical drug discovery and safety pharmacology models that closely mimic the physiological human environment[37]. In particular, the availability of hiPSC models now makes it possible to generate human cardiomyocytes *in vitro* from both healthy individuals and from patients with cardiac abnormalities[38]. A wide range of scientific studies has shown how *in vitro* hiPSC-cardiomyocyte models closely mimic their *in vivo* disease phenotype[39-41]. This allows the pharmaceutical industry to design and develop compounds on disease-specific human *in vitro* models[38]. Minimal qualifications that are necessary for hiPSC-CMs to be of interest for pharmaceutical industry is the high quality of cells, high purity, and high yields; all contributing to the high predictability and cost-effectiveness of the screening models to be developed. Advanced protocols to yield high numbers of cardiomyocytes are present and constantly subject to improvements[10, 11]. Furthermore, current issues that are being studied, which may be crucial for reliable prediction of drug efficacy and/or toxicity, are the immature nature of these cardiomyocytes[42], the lack of cardiomyocyte subtype purity[43], and the physiological conditions of cardiomyocyte models that may not reflect the *in vivo* adult heart.

The combinatorial use of fluorescent reporters and high-content analysis in high-through-put screening assays

For cost-effective drug discovery, safety and toxicity screening assays, high-throughput screenings and high content analysis are required. Detailed opportunities for fluorescent reporters in such assays are described in **Chapter 2**. In **Chapter 7** we show how a live fluorescent proliferation reporter in PSCs may be used to detect cardiac progenitor proliferation, available for high-content detection in a cardiomyocyte monolayer differentiation assay. Additionally, such an assay could be useful for studying cardiomyocyte proliferation, which could for example be of interest for therapies of the damaged heart. To note, the development of assay-specific automated imaging analysis software is a prerequisite for high-content screenings. The combinatorial use of such technologies may together be a powerful opportunity to develop biochemical and cell-based assays for compound efficacy and/or toxicity screenings.

Future physiological relevant cardiac assays

The development of predictive physiological relevant *in vitro* cardiac assays for drug discovery and safety pharmacology is a whole new era to be explored in the short-term, especially since legislation to decrease adverse drug reactions (pharmacovigilance) are being strengthened (Pharmacovigilance Legislation EU, effective since 2012), and new legislation is being developed to reduce the use of animal research in drug efficacy and risk management studies (Directive 2010/63/EU, effective since 2013).

Cardiotoxicity and safety pharmacology remain major concerns during drug development, with increased pro-arrhythmic potential being one of the main concerns. The pharmaceutical industry is challenged by the growing costs of research and development and cannot afford drug attrition in late phases of development or withdrawals of approved drugs[44]. hPSC-CM based assays are predicted as a new paradigm for assessment of cardiac risk in humans. Assays may include action potential analysis, contraction rates and features, specific ion-current analysis, cell viability measurements, force of contraction measurements, sarcomeric organization, metabolic activity, cell size (e.g. hypertrophy), or activation of specific signalling pathways. Moreover, current protocols are mostly developed in a 2-dimensional matter. It may be of interest to develop complex 3-dimensional models, mimicking *in vivo* environment, consisting of a combination of multiple cell-types, such as endothelial cells, smooth muscle cells, fibroblasts, and cardiomyocyte subtypes. Also physical contraction and the presence of a haemodynamic blood flow may be of interest to develop in future assays. In addition, disease modelling and drug efficacy and toxicity may also be studied on organ-system level, in the presence of a blood flow system, with the interplay of important organs, including liver, heart, and kidneys.

The FDA CiPA Initiative

One of the factors that currently drives the potential use of hPSC-CMs in drug development, is the FDA (U.S. Food and Drug Administration) initiative for a Comprehensive In Vitro Pro-arrhythmia Assay (CIPA). Its proposal is to increase the understanding of the electrophysiological mechanism behind pro-arrhythmic risk of drugs by the use of hPSC-CMs and *in silico* prediction models[45-48]. hPSC-CM based assays should be a new paradigm for assessment of clinical potential of Torsades de Pointes (TdP), a specific type of abnormal heart rhythm that can potentially lead to sudden cardiac death. Recent withdrawals of prescription drugs from clinical use because of unexpected side effects on the heart have highlighted the need for more reliable cardiac safety pharmacology assays. Block of the human Ether-a-go go Related Gene (hERG) ion channel in particular is associated with life-threatening arrhythmias[45], but assaying concomitant block of multiple ion channels appears to improve predictability of TdP arrhythmia, indicating the involvement of more ion currents[46]. For standardization, hPSC-CM assays should show a higher specificity and give less “false positives” than the current hERG assay (hERG channel overexpression in HEK293 cells), leading to more sophisticated ECG modeling of drug effects in the future.

CONCLUDING REMARKS AND PERSPECTIVES

In this thesis, we aimed to decipher the molecular mechanisms important for early cardiac lineage commitment and further differentiation through the generation of human PSC fluorescent reporter lines. Here, the activity of cardiac progenitor and/or cardiomyocyte subtype specific genes was visualized and followed through fluorescent reporter proteins.

Knowledge of signals involved in cardiomyocyte differentiation from PSCs is growing rapidly, and efficient protocols are being developed. However, an in-depth understanding of key cardiac regulators, at epigenetic and genetic level, involved in cardiac (sub) lineage commitment is still sparse. The advanced development of new technologies and efficient differentiation protocols will allow us to discover how complex networks interrogate to form the developing human heart, composed of sublineages of cardiac cells. Cardiac developmental biology is the basis for understanding the onset of congenital heart disease, and the development of predictive complex physiologically relevant cell-based assays for disease modelling, drug discovery and pharmacokinetics. In the end, the development of animal-replacing and highly predictive hPSC-derived cell-based assays is under high attention and promise, and currently underway.

REFERENCES

1. Ovchinnikov DA, Hidalgo A, Yang S-K, et al. Isolation of Contractile Cardiomyocytes from Human Pluripotent Stem-Cell-Derived Cardiomyogenic Cultures Using a Human NCX1-EGFPReporter. STEM CELLS AND DEVELOPMENT 2015;24(1):11–20.
2. Kita-Matsuo H, Barcova M, Prigozhina N, et al. Lentiviral Vectors and Protocols for Creation of Stable hESC Lines for Fluorescent Tracking and Drug Resistance Selection of Cardiomyocytes. PLOS ONE 2009;4(4):e5046.
3. 3 Elliott DA, Braam SR, Koutsis K, et al. NKX2-5eGFP/w hESCs for isolation of human cardiac progenitors and cardiomyocytes. NAT METH 2011;8(12):1037–1040.
4. Dubois NC, Craft AM, Sharma P, et al. SIRPA is a specific cell-surface marker for isolating cardiomyocytes derived from human pluripotent stem cells. NATURE BIOTECHNOLOGY 2011;29(11):1011–1018.
5. Paige SL, Thomas S, Stoick-Cooper CL, et al. A Temporal Chromatin Signature in Human Embryonic Stem Cells Identifies Regulators of Cardiac Development. CELL 2012;151(1):221–232.
6. Beqqali A, Kloots J, Ward-van Oostwaard D, et al. Genome-Wide Transcriptional Profiling of Human Embryonic Stem Cells Differentiating to Cardiomyocytes. STEM CELLS 2006;24(8):1956–1967.
7. Berry FB, Miura Y, Mihara K, et al. Positive and negative regulation of myogenic differentiation of C2C12 cells by isoforms of the multiple homeodomain zinc finger transcription factor ATBF1. JOURNAL OF BIOLOGICAL … 2001;(276):25057–25065.
8. Wamstad JA, Alexander JM, Truty RM, et al. Dynamic and Coordinated Epigenetic Regulation of Developmental Transitions in the Cardiac Lineage. CELL 2012;151(1):206–220.
9. Denholtz M, Bonora G, Chronis C, et al. Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. CELL STEM CELL 2013;13(5):602–616.
10. Burridge PW, Matsa E, Shukla P, et al. Chemically defined generation of human cardiomyocytes. NAT METH 2014:–.
11. Lian X, Zhang J, Azarin SM, et al. Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/β-catenin signaling under fully defined conditions. NATURE PROTOCOLS 2013;8(1):162–175.
12. Willems E, Spiering S, Davidovics H, et al. Small-Molecule Inhibitors of the Wnt Pathway Potently Promote Cardiomyocytes From Human Embryonic Stem Cell–Derived MesodermNovelty and Significance. CIRCULATION … 2011.
13. Chan SS-K, Shi X, Toyama A, et al. Mesp1 Patterns Mesoderm into Cardiac, Hematopoietic, or Skeletal Myogenic Progenitors in a Context-Dependent Manner. CELL STEM CELL 2013;12(5):587–601.
14. Cai W, Albini S, Wei K, et al. Coordinate Nodal and BMP inhibition directs Baf60c-dependent cardiomyocyte commitment. GENETICS 2013;27(21):2332–2344.
15. Behfar A, Zingman LV, Hodgson DM, et al. Stem cell

differentiation requires a paracrine pathway in the heart. THE FASEB JOURNAL 2002;16(12):1558–1566.

16. Birket MJ, Mummery CL. Pluripotent stem cell derived cardiovascular progenitors--a developmental perspective. DEVELOPMENTAL BIOLOGY 2015;400(2):169–179.
17. Ardehali R, Ali SR, Inlay MA, et al. Prospective isolation of human embryonic stem cell-derived cardiovascular progenitors that integrate into human fetal heart tissue. PROCEEDINGS OF THE ... 2013;110(9):3405–3410.
18. Yang L, Soonpaa MH, Adler ED, et al. Human cardiovascular progenitor cells develop from a KDR+ embryonic-stem-cell-derived population. NATURE 2008;453(7194):524–528.
19. Kattman SJ, Witty AD, Gagliardi M, et al. Stage-Specific Optimization of Activin/Nodal and BMP Signaling Promotes Cardiac Differentiation of Mouse and Human Pluripotent Stem Cell Lines. CELL STEM CELL 2011;8(2):228–240.
20. Cao N, Liang H, Huang J, et al. Highly efficient induction and long-term maintenance of multipotent cardiovascular progenitors from human pluripotent stem cells under defined conditions. NATURE PUBLISHING GROUP 2013.
21. Birket MJ, Ribeiro MC, Verkerk AO, et al. Expansion and patterning of cardiovascular progenitors derived from human pluripotent stem cells. NATURE BIOTECHNOLOGY 2015;33(9):970–979.
22. MULLER M. Selection of ventricular-like cardiomyocytes from ES cells in vitro. THE FASEB JOURNAL 2000;14(15):2540–2548.
23. Bondue A, Tannler S, Chiapparo G, et al. Defining the earliest step of cardiovascular progenitor specification during embryonic stem cell differentiation. THE JOURNAL OF CELL BIOLOGY 2011;192(5):751–765.
24. Zhu W-Z, Xie Y, Moyes KW, et al. Neuregulin/ErbB Signaling Regulates Cardiac Subtype Specification in Differentiating Human Embryonic Stem CellsNovelty and Significance. CIRCULATION ... 2010;(107):776–786.
25. Davis DL, Edwards AV, Juraszek AL, et al. A GATA-6 gene heart-region-specific enhancer provides a novel means to mark and probe a discrete component of the mouse cardiac conduction system. MECHANISMS OF DEVELOPMENT 2001;108(1-2):105–119.
26. Huber I, Itzhaki I, Caspi O, et al. Identification and selection of cardiomyocytes during human embryonic stem cell differentiation. THE FASEB JOURNAL 2007;21(10):2551–2563.
27. Ribeiro MC, Tertoolen LG, Guadix JA, et al. Functional maturation of human pluripotent stem cell derived cardiomyocytes in vitro--correlation between contraction force and electrophysiology. BIOMATERIALS 2015;51:138–150.
28. Zhang Q, Jiang J, Han P, et al. Direct differentiation of atrial and ventricular myocytes from human embryonic stem cells by alternating retinoid signals. NATURE PUBLISHING GROUP 2010;21(4):579–587.
29. Devalla HD, Schwach V, Ford JW, et al. Atrial-like cardiomyocytes from human pluripotent stem cells are a robust preclinical model for assessing atrial-selective pharmacology. EMBO MOLECULAR MEDICINE 2015;7(4):394–410.
30. Josowitz R, Lu J, Falce C, et al. Identification and purification of human induced pluripotent stem cell-derived atrial-like cardiomyocytes based on sarcolipin expression. 2014.
31. Minamisawa S, Uemura N, Ishikawa Y, et al. Sarcolipin, a Novel Calcium Regulator, Is Preferentially Expressedin the Atrium and Is Downregulated in the Hypertrophic Heart. CIRCULATION JOURNAL : OFFICIAL JOURNAL OF THE JAPANESE CIRCULATION SOCIETY 2003;67:211.
32. Shaner NC, Steinbach PA, Tsien RY. A guide to choosing fluorescent proteins. NAT METH 2005;2(12):905–909.
33. Trichas G, Begbie J, Srinivas S. Use of the viral 2A peptide for bicistronic expression in transgenic mice. BMC BIOLOGY 2008;6(1):40.
34. Costa M, Dottori M, Ng E, et al. The hESC line Envy expresses high levels of GFP in all differentiated progeny. NAT METH 2005;2(4):259–260.
35. Irion S, Luche H, Gadue P, et al. Identification and targeting of the ROSA26 locus in human embryonic stem cells. NATURE BIOTECHNOLOGY 2007;25(12):1477–1482.
36. Hockemeyer D, Soldner F, Beard C, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. NATURE BIOTECHNOLOGY 2009;27(9):851–857.
37. Doherty KR, Wappel RL, Talbert DR, et al. Multi-parameter in vitro toxicity testing of crizotinib, sunitinib, erlotinib, and nilotinib in human cardiomyocytes. TOXICOLOGY AND APPLIED PHARMACOLOGY 2013;272(1):245–255.
38. Davis RP, van den Berg CW, Casini S, et al. Pluripotent stem cell models of cardiac disease and their implication for drug discovery and development. TRENDS MOL MED 2011;17(9):475–484.
39. Bellin M, Casini S, Davis RP, et al. Isogenic human pluripotent stem cell pairs reveal the role of a KCNH2 mutation in long-QT syndrome. EMBO J. 2013;32(24):3161–3175.
40. Moretti A, Bellin M, Welling A, et al. Patient-specific induced pluripotent stem-cell models for long-QT syndrome. N ENGL J MED 2010;363(15):1397–1409.
41. Matsa E, Burridge PW, Wu JC. Human Stem Cells for Modeling Heart Disease and for Drug Discovery. SCIENCE TRANSLATIONAL MEDICINE 2014;6(239):239ps6–239ps6.
42. Yang X, Pabon L, Murry CE. Engineering adolescence: maturation of human pluripotent stem cell-derived cardiomyocytes. CIRCULATION RESEARCH 2014;114(3):511–523.
43. Blazeski A, Zhu R, Hunter DW, et al. Cardiomyocytes derived from human induced pluripotent stem cells as models for normal and diseased cardiac electrophysiology and contractility. PROGRESS IN BIOPHYSICS AND MOLECULAR BIOLOGY 2012;110(2-3):166–177.
44. Polak S, Pugsley MK, Stockbridge N, et al. Early Drug Discovery Prediction of Proarrhythmia Potential and Its Covariates. AAPS J 2015;17(4):1025–1032.
45. Braam SR, Tertoolen L, van de Stolpe A, et al. Prediction of drug-induced cardiotoxicity using human embryonic stem cell-derived cardiomyocytes. STEM CELL RESEARCH 2010;4(2):107–116.
46. Kramer J, Obejero-Paz CA, Myatt G, et al. MICE models: superior to the HERG model in predicting Torsade de Pointes. SCI REP 2013;3:2100.
47. Beattie KA, Luscombe C, Williams G, et al. Evaluation of an in silico cardiac safety assay: using ion channel screening data to predict QT interval changes in the rabbit ventricular wedge. J PHARMACOL TOXICOL METHODS 2013;68(1):88–96.
48. Mirams GR, Cui Y, Sher A, et al. Simulation of multiple ion channel block provides improved early prediction of compounds' clinical torsadogenic risk. CARDIOVASCULAR RESEARCH 2011;91(1):53–61.

SUMMARY

The establishment of human pluripotent stem cell (hPSC) technology and advanced cardiomyocyte differentiation protocols has opened a new platform for regenerative medicine, cardiac development, and assay development for tissue engineering, disease modelling, and drug toxicity testing. The need for these assays will be largely increasing in our nearby future, due to strengthened legislations, aiming at decreasing adverse drug reactions, and reducing the use of animal research in drug efficacy and risk management studies. Moreover, since the ability to generate patient- and disease- specific hiPSC lines, our health care system will slowly move towards the idea of personalized medicine, a model that aims at patient-specific medical treatments, based on genetic and biomarker information. Such a model can both lower costs and increase the quality of health care, and can therefore be of high importance for a sustainable health care system in our future.

In this PhD research

For pharmaceutical applications, it is important that cardiomyocyte-based assays are of highly predictive value. The use of cardiomyocyte subtypes in such assays could be of great importance, which may include atrial or ventricular cardiomyocytes, or cells of the cardiac conduction system. Moreover, the combination of different cardiomyocyte subtypes could also play a role in successful cellular transplantation therapies of the damaged heart. Therefore, it is of great interest to develop differentiation protocols that direct towards the generation of specific subtypes of cardiomyocytes. In this thesis, we describe the generation of fluorescent stem cell reporter lines to obtain insights into molecular mechanisms that play role in the formation of early cardiac progenitors and their commitment to cardiomyocyte subtypes.

A comprehensive overview of existing fluorescent reporter lines and their developmental and translational applicability is given in **Chapter 2**. Further, we describe in **Chapter 3** how the generation of a dual cardiac reporter MESP1$^{mCherry/w}$ NKX2-5$^{eGFP/w}$ hESC line allows us to study the formation of MESP1 expressing pre-cardiac mesoderm progenitors and their further commitment towards the cardiac lineage, in which key cardiac transcription factor NKX2-5 becomes expressed. We show the characteristics of MESP1-expressing progenitors and their cardiac derivatives based on gene expression profiling and surface marker expression analysis **(Chapter 3,4)**. In **Chapter 4**, we point towards putative cardiac (co)-regulators, based on their temporal expression profile upon cardiac differentiation, and their predicted protein-protein interaction with established key cardiac transcription factors. Functional experiments are required to identify their definite role.

In Chapter 5 and 6 we describe the generation of dual cardiac reporters TBX3e-mCherry-NKX2-5$^{eGFP/w}$ and MYL2-T2A-mCherry-NKX2-5$^{eGFP/w}$ hESC lines, aiming at obtaining either human nodal-like or ventricular cardiomyocytes *in vitro*, respectively. We discuss the technical and biological hurdles that we faced upon developing directed differentiaton protcols towards nodal-like cardiomyocytes. Further, we showed how MYL2 expression levels in early ventricular cardiomyocytes were low, and did only increase after long-term culture. Interestingly, we showed how MYL2 levels could be induced upon culture in maturation-inducing medium, which could indicate MYL2 as potential cardiomyocyte maturation marker.

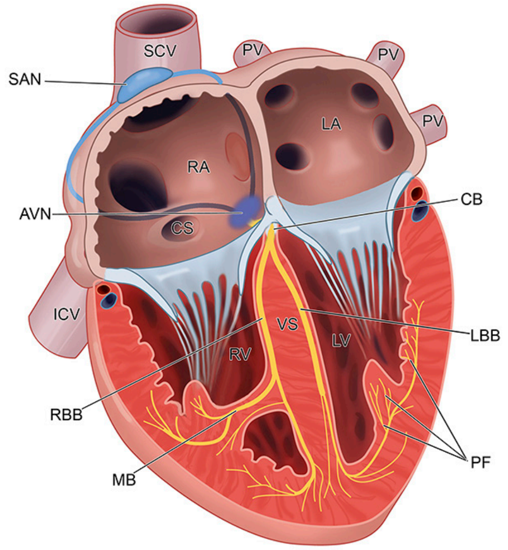
Further, in **Chapter 7**, we showed how proliferation reporter Anillin, fused to fluorescent protein eGFP, could be used to visualize distinct phases of the cell cyle, including cell division. We indicate how the visualization of proliferation using such a marker, in combination with automated quantification software, could be of potential use for high-throughput screenings of molecules playing a role in cardiac progenitor development and/or cardiomyocyte regeneration.

In conclusion **(Chapter 8)**, the innovation of advanced molecular technologies and efficient directed cardiomyocyte differentiation protocols will contribute to a better understanding of the complexicity of how epigenetics and transcriptional networks play a crucial role in cardiac lineage commitment and the onset of disease. In-depth knowledge into cardiomyocyte differentiation and development is a prerequisite for translational applications of cardiomyocytes, such as cellular therapy, or drug discovery/safety screening assays.

SAMENVATTING

Hoe zit een hart in elkaar?

Het hart is het eerst werkende orgaan tijdens de ontwikkeling van zoogdieren, zoals de mens. Al vlak na de vorming brengt het met pulserende bewegingen de bloedstroom op gang, wat vervolgens zorgt voor de zuurstof- en voedingsstof voorziening van het gehele embryo. Een volwassen hart is complex en bestaat uit vier afzonderlijke hartkamers: twee hartboezems (*de atria*), en twee hartkamers (*de ventrikels*) **(Figuur 1)**. Zuurstofrijk bloed stroomt vanuit de longen via het linkerboezem naar het linker ventrikel. Door de samentrekking van de dikke gespierde ventrikelwand, wordt dit bloed het lichaam in gepompt om de organen van zuurstof te voorzien. Zuurstofarm bloed komt vervolgens vanuit het lichaam in het rechterboezem terecht, en wordt via het rechterventrikel de longen in gepompt, zodat het bloed weer voorzien kan worden van zuurstof en terug naar het linkerboezem kan stromen. Het hart heeft een gespecialiseerd geleidingssysteem om te zorgen dat de verschillende kamers in de juiste volgorde samentrekken. De pacemakercellen van het hart geven de eerste prikkel en zorgen dat de twee boezems samentrekken. Dit signaal komt vervolgens binnen in de atrioventriculaire knoop (AV-knoop) en wordt na een korte vertraging doorgestuurd naar een bundel van vezels, die vertakt, en via Purkinje fibers verbonden is met de hartspiercellen van de ventrikels, die dan samentrekken.



Figuur 1. Het volwassen hart met het bijbehorende geleidingssysteem. SAN: sinoatrial node. AVN: Atrioventricular node. CB: Common bundle/His Bundle. RA: right atrium. SCV: superior caval vein. LBB/RBB: left/right bundle branches. PF: Purkinje Fiber network. CS: coronary sinus. IVC: inferior vena cava. LV: left ventricule. MB: moderator band. PV: pulmonary vein. RV: right ventricle. *Jongbloed et al. Differentiation 2012.*

Waarom doen we onderzoek naar de ontwikkeling van het hart?

De vorming en ontwikkeling van het hart is complex. Genetische afwijkingen spelen vaak al tijdens de vroege ontwikkeling een grote onderliggende rol in het ontstaan van hartziekten. Echter, soms komen afwijkingen pas tot uiting in een later stadium van het leven, zoals bij hevige inspanning of ouderdom. Daarnaast wordt het ontstaan van ziekten beïnvloed door omgevingsfactoren, zoals het gebruik van medicijnen. Om het ontstaan van hartziekten te kunnen voorkomen, en te genezen met de juiste medicijnen, is het van belang een gedetailleerd begrip te hebben van de ontwikkeling van het menselijk (*humane*) hart, en de reactie van het hart op specifieke fysiologische omstandigheden. Het gebruik van humane (ziekte)modellen is hiervoor een pré.

Kloppende hartcellen kweken in een schaaltje

Omdat het gebruik van humane embryo's beperkt is, en in veel landen ethisch omstreden, worden vooral proefdieren en eenvoudige cel modellen gebruikt om deze vraagstukken te beantwoorden. Echter, vaak blijkt de stap tussen mens en dier groot. Helaas liggen goede humane onderzoeksmodellen niet voor het oprapen. Het kweken van humane hartcellen (*cardiomyocyten*) uit stamcellen biedt daarom een exclusieve oplossing.

Vanuit humane pluripotente stamcellen (hPSCs) zijn we tegenwoordig in staat grote hoeveelheden cardiomyocyten te kweken. Onder hPSCs verstaan we embryonale stamcellen (hESCs) en geïnduceerde pluripotente stamcellen (hiPSCs). hESCs worden, sinds 1998, geïsoleerd uit een celklompje van een gerijpte bevruchte eicel, blastocyst genaamd, wat een overblijfsel is van IVF behandelingen. HiPSCs worden, sinds 2007, verkregen uit menig celtype van een volwassen mens door deze te "induceren" met een specifieke combinatie van genen, en zijn daarom ethisch minder omstreden. Uit hPSCs kunnen we bijna elke cel van het menselijk lichaam vormen door ze in een gecontroleerde omgeving een specifieke richting op te sturen met behulp van groeifactoren of chemische moleculen.

Het kweken van cardiomyocyten uit hPSCs geeft ons een unieke kans om de embryonale ontwikkeling van het menselijke hart in een kweekschaal te kunnen bestuderen, en daarbij inzicht te krijgen in het ontstaan van aangeboren hartziekten.

Klinische toepassingen van in vitro gekweekte hartcellen

Sinds we in staat zijn hiPSCs te creëren, bieden deze *in vitro* verkregen hartcellen ook een groot potentieel voor toekomstige klinische en farmaceutische toepassingen, zoals celtransplantaties ter vervanging van beschadigd hartweefsel, en het opzetten van effectiviteits/toxiciteit-screening-experimenten (*assays*) voor de ontwikkeling van nieuwe medicijnen.

De ontwikkeling van een nieuw medicijn begint bij de ontdekking van nieuwe stoffen in het "drug discovery" stadium in het laboratorium. Na uitvoerige testen op cel- en proefdiermodellen, wordt een strenge selectie van medicijnen getest in klinische studies op mensen. Echter, juist in dit stadium blijken veel medicijnen niet voor elke patiënt effectief te zijn, of voor negatieve bijwerkingen te zorgen, die in het initiële "drug discovery" stadium niet zijn opgepikt in de huidige screenings-modellen. Daarnaast zijn er momenteel veel medicijnen op de markt met negatieve bijwerkingen voor het hart, zoals veel chemotherapeutische middelen, die mogelijk in een goede toxiciteits-screening-assay uitgeselecteerd hadden kunnen worden.

De vraag naar betrouwbare assays zal daarom in de toekomst steeds groter worden, mede door de aanstelling van nieuwe wetgeving waarbij het gebruik van proefdieren verminderd zal moeten worden, en de wetgeving rondom de gevaren voor negatieve bijwerkingen van medicijnen steeds strenger wordt. Daarnaast zal de medische wereld zich, dankzij de ontwikkeling van patiënt-specifieke iPS modellen, steeds meer gaan richten op patiënt-specifieke medicatie (*personalized medicine*), wat zal kunnen leiden tot een hogere kwaliteit van zorg en een bijdrage aan de ontwikkeling van een duurzaam zorgsysteem, wat van groot toekomstig belang zal zijn in een vergrijzende maatschappij.

Dit promotieonderzoek

Voor farmaceutische toepassingen is het van belang dat de voorspellende waarde van cardiomyocyte assays zo betrouwbaar mogelijk is, waarbij het gebruik van specifieke cardiomyocyte subtypes van invloed kan zijn. Voorbeelden van subtypes zijn: hartkamer-(ventrikel) of hartboezem-(atrium) cellen, of cellen van het geleidingssysteem, zoals pacemakercellen **(Figuur 1)**. Daarnaast is gebleken uit cel transplantatie studies in dieren, dat de combinatie van de juiste celtypes van groot belang is voor vervanging van beschadigd hartweefsel. Om deze redenen is het voor zowel de academische als de farmaceutische wereld interessant om uit humane pluripotente stamcellen een zuivere populatie van specifieke subtypes hartcellen te verkrijgen. In deze thesis beschrijven we de generatie van fluorescente stamcel reporter lijnen als methode om inzicht te verkrijgen in de moleculaire signalen die een rol spelen bij de vorming van specifieke voorlopercellen en cardiomyocyte subtypes.

Hoofdstuk 1 geeft een korte inleiding in de mogelijkheden van het gebruik van pluripotente stamcellen en translationele toepassingen. We geven aan hoe kennis, verkregen uit embryonale diermodellen zoals de muis, heeft geleid tot de ontwikkeling van efficiënte *in vitro* kweekprotocollen voor het verkrijgen (*differentiëren*) van cardiomyocyten uit humane pluripotente stamcellen. We beschrijven hoe genetische netwerken* *in vivo*, op een vergelijkbare manier van belang zijn tijdens de vorming van cardiomyocyten *in vitro*, wat aangeeft hoe nauw onze in vitro kweekmodellen overeenkomen met in vivo ontwikkelingsbiologie.

*: genetische netwerken bestaan uit een netwerk van eiwitten (de producten van genen), die door complexe interacties met elkaar en met het DNA, de activiteit van andere genen kunnen beïnvloeden.

Hoofdstuk 2 geeft een uitgebreid overzicht van de tot nu toe beschreven *in vitro* fluorescente pluripotente stamcel reporter lijnen* die gebruikt zijn voor het visualiseren en verkrijgen van *in vitro* hartcellen. Daarnaast beschrijven we de potentie van fluorescente reportermodellen voor de ontwikkeling van assays op het gebied van medicijnontwikkeling, toxiciteitscreenings, en ziektemodellen, en laten we zien hoe fluorescente reporters gebruikt kunnen worden in transplantatiestudies van het hart.

*: in een fluorescente stamcel lijn is het DNA van een stamcel gemodificeerd; we vervangen de genetische code van een gen dat we willen volgen, door de genetische codering voor een fluorescent eiwit. Wanneer het specifieke gen actief wordt in een specifiek celtype, zal het fluorescente eiwit worden aangemaakt en zichtbaar worden in een cel.

Hoofdstuk 3 beschrijft de opbouw van een duale fluorescente reporter lijn: een MESP1$^{mCherry/w}$NKX2-5$^{eGFP/w}$ humane embryonale stamcel lijn. Deze lijn maakt het mogelijk om tijdens *in vitro* cardiomyocyte differentiatie, de ontwikkeling van vroege voorlopercellen van het hart (en het gen MESP1 tot expressie brengen, nog voordat het hart is gevormd) te volgen, om vervolgens te bestuderen welke moleculaire mechanismen een rol spelen bij hun verdere differentiatie naar cardiomyocyten (en het gen NKX2-5 tot expressie brengen, wat aanwezig is kort voor de vorming, en in het gehele hart). De expressie van bovengenoemde genen in specifieke celtypes kan gevolgd worden doordat ze gekoppeld zijn aan de genetische codes voor fluorescente eiwitten, zoals mCherry of eGFP. Deze eiwitten worden bij een specifieke lichtlengte aangeslagen (excitatie), en zenden vervolgens een eiwit-specifiek lichtpatroon uit (emissie).
In dit hoofdstuk isoleren en karakteriseren we MESP1-mCherry positieve voorloperellen op basis van de aanwezigheid van het mCherry eiwit. We kijken naar de expressie van celmembraan markers en naar het algehele genexpressie patroon. We identificeren specifieke verschillen met eerdere *in vitro* studies in muis, en geven daarmee het belang aan van *in vitro* studies in menselijke celtypes.

In **Hoofdstuk 4** differentiëren we geïsoleerde MESP1 voorlopercellen naar NKX2-5 positieve hartcellen en meten we genexpressie profielen op verschillende tijdstippen, waaronder dag 5, dag 7, dag 10 en dag 14. Op basis van temporale genexpressie patronen, voorspellende eiwit-eiwit interacties met huidige belangrijke transcriptionele genen, en hun DNA bindende eigenschappen, identificeren we potentiele (co-)regulators van cardiomyocyte differentiatie. Functionele experimenten zullen moeten uitwijzen wat de rol van deze potentiele regulators zal zijn.

In **Hoofdstuk 5** beschrijven we de opbouw van een duale TBX3e-mCherry-NKX2-5$^{eGFP/w}$ hESC reporter line. TBX3 komt in het hart specifiek tot expressie in het geleidingssysteem. Voor studies naar afwijkingen binnen dit systeem is het van belang om de moleculaire mechanismen die een rol spelen bij het ontstaan van conductiesysteem cardiomyocyten te begrijpen. Met behulp van deze duale reporter lijn hoopten we TBX3 positieve cardiomyocyten te verkrijgen voor verdere karakterisatie. Echter, het blijkt moeilijk te zijn om met de huidige kennis van het ontwikkelende geleidingssystem in vivo, protocollen op te zetten om *in vitro* pluripotente stamcellen te differentiëren naar deze celtypes. Ook technische uitdagen kunnen bij de afwezigheid van TBX3 positieve cardiomyocyten een rol hebben gespeeld.

Hoofdstuk 6 beschrijft de opbouw van een derde duale cardiale reporter: MYL2-T2A-mCherry-NKX2-5$^{eGFP/w}$ hESCs. Deze reporter lijn maakt het mogelijk om specifiek ventriculaire cardiomyocyten te isoleren, door middel van co-expressie van MYL2 en NKX2-5. In deze studie laten we zien dat de ventriculaire marker MYL2 pas laat tijdens differentiatie hoog tot expressie komt, en daarom mogelijk een graadmeter is voor cardiomyocyte maturatie. Daarnaast laten we zien dat maturatie-gerelateerde factoren als dexamethason, IGF-1, en thyroid hormoon MYL2 expressie in cardiomyocyten verhogen.

In **Hoofdstuk 7** bestuderen we het delingsvermogen (proliferatie capaciteit) van MESP1 positieve voorlopercellen van het hart. Het verkrijgen van een oneindige voorraad van voorlopercellen kan van grote interesse zijn voor transplantatiestudies of voor grootschalige experimentele assays. We laten met time-lapse imaging zien hoe we de proliferatie van deze voorlopercellen kunnen visualiseren met behulp van het eiwit Aniline, wat in deze assay gekoppeld is aan het fluorescente eiwit eGFP. Daarnaast laten we zien hoe de ontwikkeling van automatische kwantificatie software, in combinatie met deze fluorescente proliferatie reporter, mogelijk interessant kan zijn voor grootschalige proliferatie-inducerende screening assays op hartvoorlopercellen of cardiomyocyten.

Hoofdstuk 8 omvat de algemene discussie van dit proefschrift. We concluderen dat de ontwikkeling van nieuwe moleculaire technieken en efficiënte differentiatie protocollen gaat bijdragen aan een beter begrip van de complexiteit van epigenetische en genetische netwerken die cruciaal zijn voor de embryonale ontwikkeling van het humane hart, bestaande uit een groot aantal subtypes hartcellen. Kennis van humane hartontwikkeling is de basis voor het verkrijgen van inzicht in het ontstaan van aangeboren en verkregen hartafwijkingen, en de ontwikkeling van voorspellende complexe fysiologisch-relevante cardiomyocyte-assays voor menselijke ziektemodellen, medicijn ontwikkeling- en testsystemen. De ontwikkeling van proefdiervrije en hoog-voorspellende hPSC-gebaseerde cardiomyocyte assays is veelbelovend.

CURRICULUM VITAE

Sabine Charlotte Den Hartogh was born on July 16th, 1987 in Nijkerk, a small village on the Veluwe, the Netherlands. In 2005, she graduated at Meerwegen College Farel in Amersfoort. In that same year, she started her studies Biomedical Sciences at the Graduate School of Life Sciences, Utrecht University.

She was introduced to the field of stem cell biology for the heart when she carried out an internship in the laboratory of Prof. dr. Deepak Srivastava at the Gladstone Institute in San Francisco, where she studied the role of MicroRNA-24 in cardiomyocyte survival upon transplantation in the heart, under supervision of dr. Li Qian and dr. Linda van Laake. Then, she returned to Utrecht where she wrote a review on cellular therapy of the heart, in collaboration with dr. Linda Van Laake at the UMC Utrecht, to finally obtain her Master of Science degree in April 2011.

In May 2011, she joined the lab of Dr. Robert Passier as a PhD-candidate, at the department of Anatomy and Embryology of Leiden University Medical Centre, headed by Prof. dr. Christine Mummery. Here, she published several studies and a review on the generation of human embryonic stem cell reporter lines in order to understand cardiac development and to optimize cell culture conditions for directed differentiation towards cardiomyocyte (progenitor) subtypes for pharmaceutical applications and/or regenerative medicine.

Since October 2015, she is applying her in-depth knowledge of cardiovascular biology in a translational matter by starting on a position at Pluriomics, a biotechnology start-up company in Leiden Bioscience Park, focusing on the development of cardiovascular cell-based assays for cardiac safety pharmacology and drug discovery.

LIST OF PUBLICATIONS

Cao L, **den Hartogh SC**, Verbeek FJ, Passier R. Automatic image analysis for distinction of cell division based on eGFP-Anillin signal in human MESP1 progenitors. *Submitted.*

Den Hartogh SC, Wolstencroft K, Mummery CL, Passier R. A comprehensive gene expression analysis at sequential stages of in vitro cardiac differentiation from isolated MESP1-expressing-mesoderm progenitors. *Scientific Reports* ***6****, 2016. doi:10.1038/srep19386*

Den Hartogh SC, Passier R. Concise Review: Fluorescent reporters in human pluripotent stem cells: contributions to cardiac differentiation and their applications in cardiac disease and toxicity. *Stem Cells. 2016 Jan;34(1):13-26. doi: 10.1002/stem.21962015.*

Birket MJ, Ribeiro MC, Verkerk AO, Ward D, Leitoguinho AR, **Den Hartogh SC**, Orlova VV, Devalla HD, Schwach V, Bellin M, Passier R, Mummery CL. Expansion and patterning of cardiovascular progenitors derived from human pluripotent stem cells. *Nat Biotechnol. 2015 Jul 20.*

Den Hartogh SC, Schreurs C, Monshouwer-Kloots JJ, Davis RP, Elliott DA, Mummery CL, Passier R. Dual reporter MESP1 mCherry/w-NKX2-5 eGFP/w hESCs enable studying early human cardiac differentiation. *Stem Cells. 2015 Jan;33(1):56-67*

Hartogh Den SC, Sluijter J, Van Laake LW, et al. Chapter 11. Cellular therapy for the infarcted myocardium. *In: Translational Cardiology: Molecular Basis of Cardiac Metabolism, Cardiac Remodeling, Translational Therapies, and Imaging Techniques. Cam Patterson, Monte Villis. Springer; 2012.*

SCIENTIFIC AWARDS

Recipient van Wijck-Stam-Caspers award, 2010