State of the heart : the promise of pluripotent stem cell-derived cardiomyocytes in disease modelling, differentiation and development

Berg, C.W. van den

CHAPTER 5

Comparison of genetically marked NKX2-5-eGFP human induced pluripotent- and embryonic stem cells shows extensive molecular and functional similarity during early but not late cardiomyogenesis

Cathelijne W. van den Berg [1], Satoshi Okawa [2], Simona Casini [1], Antonio del Sol [2], David A. Elliott [3], Christine L. Mummery [1], Richard P. Davis [1]

In preparation

1 Department of Anatomy and Embryology, Leiden University Medical Center, Leiden, The Netherlands
2 Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Luxembourg
3 Murdoch Childrens Research Institute, The Royal Children's Hospital, Parkville, Victoria, Australia

Abstract

Human embryonic stem cells (hESCs) and induced pluripotent stem cells (hiPSCs) offer opportunities to study early human development, model genetic diseases, as well as for drug discovery and toxicity testing. While several studies have examined undifferentiated ESCs and iPSCs for transcriptional and epigenetic differences, few have compared the differentiated cell types.

In considering the suitability of hiPSCs as a model for early human cardiac development, we compared their functional and molecular characteristics with hESC-derived cardiomyocytes, a more widely studied model. To ensure cells were compared at the same development stage we targeted a hiPSC line with *eGFP* into the locus of *NKX2-5* (NKX2-5-eGFP) by homologous recombination. *NKX2-5* is a transcription factor expressed in cardiac progenitors which we have previously targeted in hESCs. Using defined culture conditions that first supported hiPSC and hESC in an undifferentiated state and allowed both to be differentiated to cardiomyocytes under identical conditions, we generated cultures in which more than 50 % of the cells expressed NKX2-5-eGFP.

The transcriptional profile of NKX2-5-eGFP expressing hiPSCs and hESCs differed little at early stages of differentiation although diverged later. There were no differences in action potentials of the cardiomyocytes derived from either hiPSCs or hESCs, demonstrating that NKX2-5-eGFP hiPSCs could replace hESCs in studying human cardiovascular differentiation.

Introduction

Human embryonic stem cells (hESCs) and induced pluripotent stem cells (hiPSCs) can differentiate into many cell types including cardiomyocytes, the contractile cells of the heart. While undifferentiated hiPSCs and hESCs are similar in morphology, expression of pluripotency markers, capacity for self-renewal and global gene expression profiles, differences have been described in their epigenetic state and transcriptome profiles [1]. It is unclear whether this impacts their capacity for differentiation or function of their somatic derivatives since few direct comparisons have been carried out, although at least two studies have described variability in the differentiation potential of hiPSC and hESC to the cardiac lineage [2,3]. The most appropriate way of comparing hiPSC and hESC directly may not be time after induction of differentiation but relative to developmental events when cells are making fate decisions. For the heart, this would be when cardiac mesoderm cells commit to become cardiac progenitors, and later when contractile cardiomyocytes develop. Direct comparison of these populations has been limited by the lack of cardiac-specific surface markers on which they could be selected, especially for cardiac progenitors. Alternatives include cardiac-specific fluorescent reporter cell lines that allow isolation of cardiac-enriched populations from both hESCs and hiPSCs [4,5]. The cardiac transcription factor *NKX2-5* is critical in early heart development, marking a cardiac progenitor that subsequently becomes a more committed cardiovascular population during differentiation [4,6]. It is thus well-suited for this purpose. Sequences encoding enhanced GFP (*eGFP*) have previously been targeted into the *NKX2-5* locus of hESCs with GFP expression reflecting endogenous *NKX2-5* (NKX2-5-eGFP), enabling both quantification of cardiac differentiation and purification of (GFP-expressing) cardiac progenitor cells and cardiomyocytes [4].

Here we describe the generation of an analogous hiPSC line in which we targeted the *NKX2-5* locus with *eGFP* using the same construct as previously in hESCs. These two unique cardiac reporter lines allowed isolation of NKX2-5-eGFP$^+$ expressing cardiac populations from both hESCs and hiPSCs at specific fate decision points during their differentiation, facilitating comparative analyses at transcriptional, protein and functional levels. The results demonstrated that at early differentiation stages hESCs and hiPSCs are phenotypically very similar in their expression of specific markers and genes and in their electrophysiological characteristics. However with prolonged culture, the transcriptional phenotypes diverged.

Materials and Methods

Generation and identification of targeted *NKX2-5$^{GFP/w}$* hiPSCs

The hiPSC line, DF4-3-7T [7] (WiCell) was targeted as previously described for hESCs [8,9]. Targeted clones were identified using a PCR-based screening strategy combined with Sanger sequencing, and confirmed with Southern blot analysis. The loxP-flanked selection cassette was excised using Cre recombinase and the resulting line subcloned as described previously [10]. Karyotype analysis was performed using COBRA-FISH and 20 metaphase spreads were analysed [11].

Cell culture and differentiation

The previously generated *NKX2-5$^{eGFP/w}$* MEL1 hESC (hESC-NKX2-5) [4] and the *NKX2-5$^{eGFP/w}$* hiPSC (hiPSC-NKX2-5) lines were maintained in culture either on mouse embryonic fibroblasts (MEFs) in standard medium [12], or on vitronectin-coated surfaces (Greiner Bio One) in Essential 8 (E8) medium (Life Technologies). For differentiation, cells were plated on Matrigel (BD) in E8 medium and treated three days later with 1 % DMSO for 24–36 hours [13]. Pluripotent stem cells (PSCs) were differentiated using a directed monolayer protocol by incubating the cells for 3 days in BPEL (Bovine Serum Albumin (BSA) Polyvinylalchohol Essential Lipids) differentiation medium [14] with Activin A (20 ng/mL, R&D), Bone morphogenetic protein 4 (BMP4; 20 ng/mL, R&D) and CHIR99021 (1.5 µM, Axon Medchem). The BPEL medium was replaced and cells incubated for a further 3 days with XAV939 (5 µM, Tocris). From day 6 of differentiation, BPEL medium was changed on the cells every 3–4 days.

Flow cytometry

For cell surface and intracellular flow cytometry experiments, hESC- and hiPSC-NKX2-5 were enzymatically dissociated using 1x TrypLE Select (Life Technologies) for 5 minutes at 37 °C. Single cells were harvested and labelled for pluripotency markers TRA-1-60 (Merck Millipore), TRA-1-81 (Merck Millipore), SSEA4 (Merck Millipore), E-CAD (Thermo Fisher Scientific) and OCT4 (Santa Cruz Biotechnology). All labellings were performed with live cells except for OCT4 where the cells were fixed and permeabilized using FIX & PERM Cell Fixation & Permeabilization Kit (Invitrogen). Primary antibodies were detected with appropriate Allophycocyanin (APC)-conjugated secondary antibodies.

Differentiated cells were enzymatically dissociated using either 1x, 5x or 10x TrypLE Select (Life Technologies) for 5–20 minutes at 37 °C. The expression of GFP was monitored at multiple time points during the differentiation. For the isolation of GFP^+ and GFP^--populations, cells were sorted using a BD ARIA III flow cytometer. Only experiments where the sorted populations were >85 % pure when reanalysed, were used for downstream applications. Cells were stained for SIRPA, VCAM1 (both Miltenyi Biotec), PDGFRA (BD Biosciences) and KDR

(R&D). All samples were analysed using a MACS-Quant VYB Flow cytometer (Miltenyi Biotec) and FlowJo Software. Flow cytometry gates were set using control cells labelled with isotype control antibodies.

Gene expression analysis

Total RNA was isolated using the Nucleospin RNA or Nucleospin RNA XS (both Macherey-Nagel) according to the manufacturer's protocol. Quality and integrity of the samples was determined using Lab-on-Chip RNA 6000 Nano on the Agilent 2100 Bioanalyzer (Agilent Technologies). Illumina TotalPrep-96 RNA Amplification was applied to generate Biotin labelled cRNA of which 750 ng was hybridized onto the Illumina HumanHT-12 v4 by ServiceXS B.V. (Leiden, The Netherlands). The samples were scanned using the Illumina iScan array scanner.

Gene expression analysis

Microarray data containing the raw intensity values was normalized by variance stabilizing normalization using the vsn R package. When a gene had more than one probe, subsequent analysis was performed with the probe that had the highest variance across samples. Concordance between different samples at different time points was determined by averaging biological replicates. The differential expression analysis was carried out by a moderated *t*-test using the limma R package [15]. The *P*-value was corrected for multiple testing by the Benjamini-Hochberg method. Genes with the adjusted *P*-value smaller than 0.05 and the mean absolute log2 fold-change larger than 1.0 were considered differentially expressed. Unsupervised hierarchical clustering of all detected genes was performed using the Euclidean distance and complete linkage method. Gene Ontology Biological Process (GO.BP) was downloaded from http://www.geneontology.org. Statistical enrichment of the GO.BP categories was identified by Fisher's exact test using the differentially up- or down-regulated genes and the *P*-value was corrected by the Benjamini-Hochberg method. All detected genes were taken as the background set. The enriched categories were defined as those with the adjusted *P*-value below 0.05 and the log2 odds ratio above 1.0.

Immunohistochemistry

Cells were fixed in 2 % paraformaldehyde and permeabilized in 0.1 % Triton (Sigma) in PBS, before being blocked in 4 % swine serum in PBS and incubated overnight with primary antibodies against GFP (Life Technologies), NKX2-5 (Santa Cruz Biotechnology), α-actinin (Sigma-Aldrich) and Troponin I (Santa Cruz Biotechnology). Primary antibodies were detected with Alexa Fluor 488 (Life Technologies) or Cy3-conjugated (Jackson ImmunoResearch) antibodies and confocal images captured using a Leica SP5 confocal laser scanning microscope (Leica Microsystems).

Electrophysiology

Electrophysiological measurements were performed on single cardiomyocytes after dissociation on day 21. Spontaneously contracting cells were paced at 1 Hz. Data was collected from 2–3 independent differentiations per line. Action potentials (APs) were recorded with the perforated patch-clamp technique using an Axopatch 200B amplifier (Molecular Devices) in a modified Tyrode solution at 37 °C. APs were elicited at 1 Hz and the upstroke velocity (V_{max}), AP duration (APD) at 50 % and 90 % (ms), AP amplitude (APA) and maximum diastolic potential (MDP) were analyzed. Data from 8–10 consecutive APs were averaged.

Results

Generation and characterization of a hiPSC-NKX2-5 reporter cell line

The hiPSC-NKX2-5 reporter line was generated through homologous recombination by inserting sequences encoding *eGFP* and a loxP-flanked positive selection cassette (loxP-PGK-neo-loxP) immediately 3' to the start codon in exon 1 of *NKX2-5* to replace part of the exon (Figure 1A). PCR amplification in combination with Sanger sequencing identified correctly targeted clones (Figure 1B). After removal of the G418 resistance cassette, the structural integrity of the targeted locus and confirmation of a single integration event were verified by Southern blot analysis (Figure 1C). In addition, one hiPSC-NKX2-5 line was cloned by single cell sorting of SSEA4$^+$ cells by flow cytometry. The parental line and subclones were phenotypically indistinguishable; all further analysis was performed with the clonal hiPSCs. The hiPSC-NKX2-5 expressed the typical stem cell markers TRA-1-60, TRA-1-81, SSEA4, OCT4 and E-CAD as determined by flow cytometric analysis (Figure 1D), and had a normal karyotype (Figure 1E).

hESC-NKX2-5 and hiPSC-NKX2-5 differentiate to cardiomyocytes

We next compared the efficiency and kinetics of the hESC- and hiPSC-NKX2-5 cell differentiation into cardiac cell types in monolayer culture (Figure 2A). Cardiac-mesoderm cells were generated in completely defined medium by addition of Activin A, BMP4 and CHIR99210 [13]. Subsequent inhibition of the Wnt/β-catenin pathway by XAV939 generated NKX2-5-eGFP$^+$ cardiac progenitor cells that subsequently became committed cardiomyocytes. These clusters of spontaneously beating cells were GFP$^+$ (Figure 2B); immunofluorescence analysis confirmed only GFP$^+$ cells co-expressed endogenous NKX2-5 (Figure 2C). GFP expression was confirmed by flow cytometry at several time points during differentiation (Figure 2D). GFP$^+$ cells were first detected on day 6, the proportion of cells expressing GFP increasing over time and peaking around day 14 with more than 50 % of the cells being GFP$^+$ (Figure 2E). The GFP$^+$ cardiomyocytes also expressed cardiac markers such as Troponin I and α-actinin, and showed characteristic sarcomeric structures (Figure 2F).

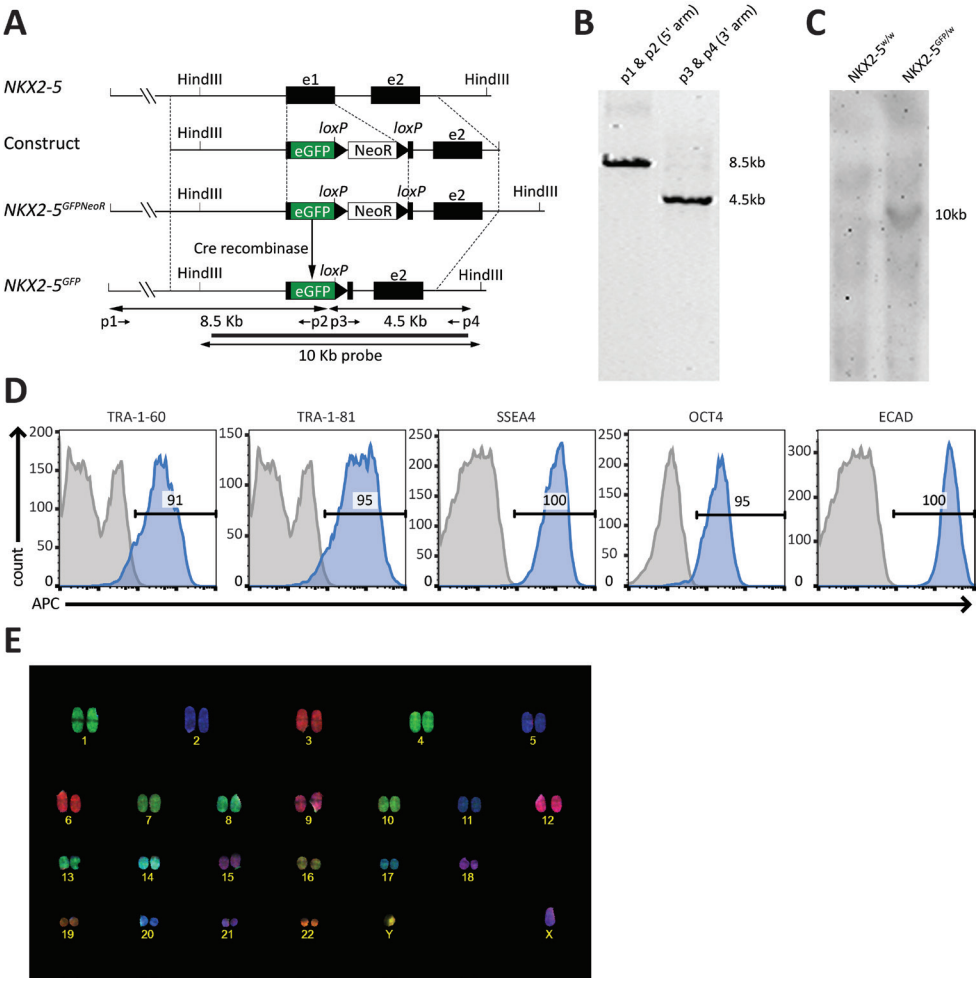

Figure 1. Characterization of the generated hiPSC-NKX2-5 line

A, Schematic of the targeting procedure to introduce *eGFP* into the first exon (e1) of *NKX2-5* by homologous recombination. Restriction sites of HINDIII, primers p1–p4 and the 10 kb Southern blot probe are displayed. **B**, Correct targeting of the gene was identified by PCR amplification. The gel shows specific bands at 8.5 kb (5' homology arm) and 4.5 kb (3' homology arm) confirming correct homologous recombination. **C**, Southern blot analysis of HINDIII-Digested genomic DNA confirmed the cells contain a single copy of the *eGFP* gene consistent with a single genetic modification at the *NKX2-5* locus. **D**, The hiPSC-NKX2-5 line expressed stem cell markers TRA-1-60, TRA-1-81, SSEA4, E-CAD and OCT4. The blue peak demonstrates the percentage of positive cells for each marker, while the grey peak marks the isotype control. **E**, The COBRA-FISH karyogram from the hiPSC-NKX2-5 line shows a normal karyotype.

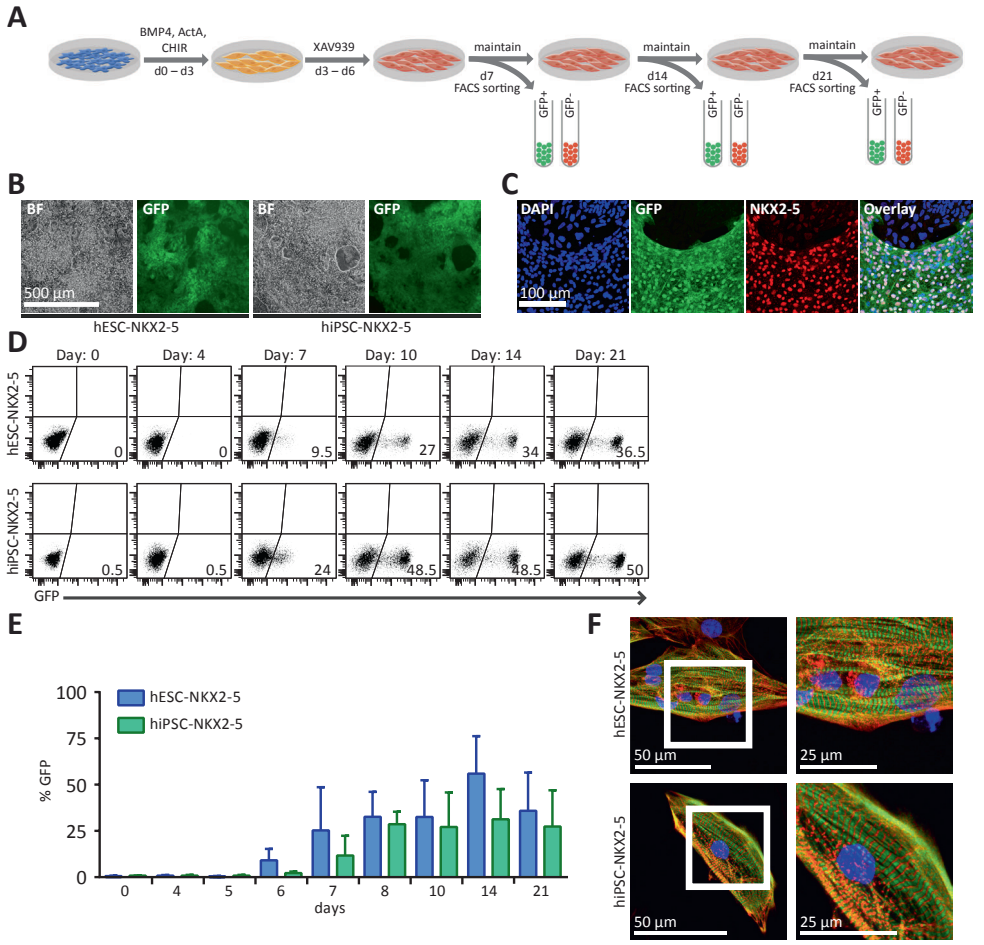

Figure 2. Differentiation of the human pluripotent stem cell (hPSC)-NKX2-5 to cardiomyocytes
A, Schematic detailing the monolayer procedure to differentiate the pluripotent stem cells to cardiomyocytes using initially Bone morphogenetic protein 4 (BMP4), Activin A (ActA), and CHIR990291, followed by XAV939 at day (d) 3. Time points for isolation of the GFP^+ and GFP^- cells by flow cytometry are indicated. **B**, In both hESC-NKX2-5 and hiPSC-NKX2-5 expression of GFP overlapped with spontaneously beating clusters. Bright field (BF) and fluorescent image are displayed. **C**, Immunofluorescence demonstrates that GFP^+ expression is restricted to (endogenous) $NKX2\text{-}5^+$ cells. Nuclei are stained in blue (DAPI). **D**, Time course of GFP expression during the monolayer differentiation procedure measured by flow cytometry. Percentages of GFP^+ cells are indicated. **E**, Expression of GFP detected by flow cytometry at multiple time points during 3 independent differentiations of both hESCs and hiPSCs. **F**, Cardiomyocytes were positive for Troponin I (green) and α-actinin (red) and showed characteristic sarcomeric structures. Nuclei are stained in blue (DAPI).

hiPSC-derived cardiomyocytes express cardiac lineage markers but are electrophysiologically immature

Cell surface proteins SIRPA and VCAM1, both markers of hESC-derived cardiomyocytes [4], were expressed in the differentiating hESC- and hiPSC-NKX2-5 lines, with VCAM1 expression first detected in a fraction SIRPA$^+$ and GFP$^+$ cells on day 7 (Figure 3A). The co-expression of SIRPA and VCAM1 was always concurrent with the expression of GFP. This increased during differentiation peaking at day 14 in both cell lines, when more than 75 % of the GFP$^+$ cells expressed both SIRPA and VCAM1. Interestingly, while the proportion of GFP$^+$ cells co-expressing these markers remained constant at later time points in the hESC-NKX2-5 line, a noticeable reduction of VCAM-expressing GFP$^+$ cells was detected consistently in the hiPSC-NKX2-5 line at day 21, suggesting an increase in the proportion of NKX2-5-eGFP$^+$ non-cardiomyocytes. As previously observed, the early NKX2-5-eGFP$^+$ cardiac mesoderm cells were positive for PDGFRA, and in both hESC- and hiPSC-NKX2-5 the proportion of PDGFRA$^+$GFP$^+$ cells decreased over time (Figure 3B). During the early stages of differentiation, a population of KDR$^+$ (kinase insert domain receptor) cells were detected that were negative for GFP. The expression of KDR decreased when the proportion of GFP$^+$ cells increased from day 7 onwards. At all time points the GFP$^+$ cells were negative for KDR (Figure 3C).

Patch-clamp electrophysiology was used to measure APs of cardiomyocytes from both cell lines (Figure 3D). Upstroke velocity (V_{max}) measured at 1 Hz was not significantly different in hESC-NKX2-5 cardiomyocytes (hESC-CMs) versus hiPSC-NKX2-5 cardiomyocytes (hiPSC-CMs): 18 V/s and 14 V/s respectively. Average data at 1 Hz for AP duration at 50 % and 90 % repolarization was also not significantly different between hESC- and hiPSC-CMs. AP amplitudes were also similar for hESC- and hiPSC-CMs (–61 and –64 mV respectively), and no differences were detected in maximal diastolic potential (103 mV for hESC-CMs and 98 mV for hiPSC-CMs) (Figure 3E).

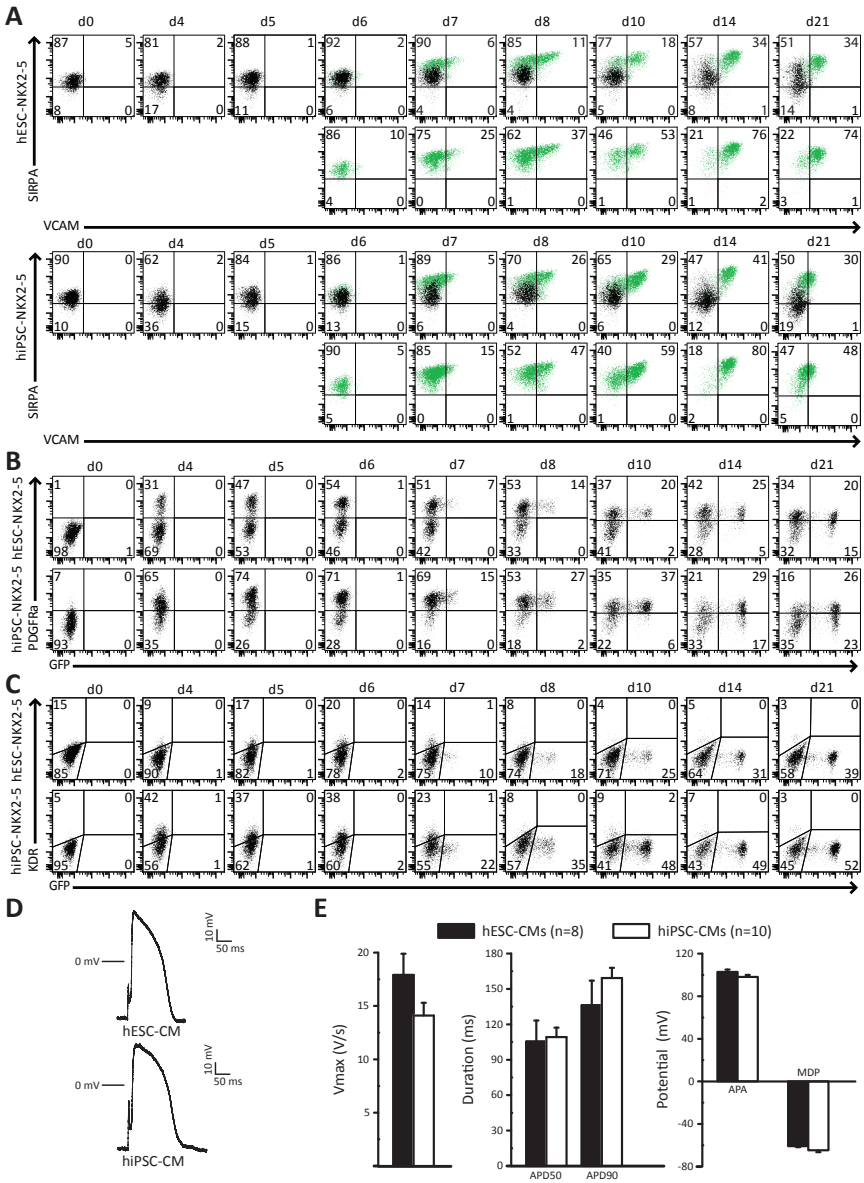

Figure 3. Cardiac differentiation kinetics and electrophysiological properties of hESC- and hiPSC-derived cardiomyocytes

A, Time course of a representative example showing temporal expression profile of SIRPA and VCAM1 in differentiating hESC- and hiPSC-NKX2-5 and their expression pattern in NKX2-5-eGFP$^+$ cells. **B** and **C**, Time course of a representative example demonstrating expression of PDGFRa (B) and KDR (C) during cardiac differentiation. **D**, Representative examples of action potential (AP) traces from hESC-NKX2-5 cardiomyocyte (hESC-CM) and hiPSC-NKX2-5 cardiomyocyte (hiPSC-CM) at 1 Hz. **E**, Average data at 1 Hz for maximal upstroke velocity (V_{max}), AP duration (APD) at 50 % and 90 % repolarization, AP amplitude (APA) and maximal diastolic potential (MDP) (Error Bars, SEM). d indicates day.

Gene expression profiles of NKX2-5-eGFP$^+$ cells from hiPSCs and hESCs are comparable at day 7 and 14 of differentiation

To investigate how similar the NKX2-5-eGFP$^+$ populations from differentiated hESC- and hiPSC-NKX2-5 were, GFP$^+$ cells were collected at 3 time points during differentiation (Figure 2A). These represented NKX2-5-eGFP$^+$ cardiac progenitor cell (CPC) population (day 7), an enriched population of NKX2-5-eGFP$^+$ cardiomyocyte population (day 14) and a divergent NKX2-5-eGFP$^+$ population (day 21). Additionally, a population of undifferentiated hESC-NKX2-5 and hiPSC-NKX2-5 (differentiation day 0) as well as the GFP$^-$ population at each of the three differentiation time points were collected. The gene expression profile of each of these populations was obtained in triplicate by microarray analysis.

Unsupervised hierarchical clustering identified three separate clusters, consisting of undifferentiated hESC-NKX2-5 and hiPSC-NKX2-5, and either GFP$^+$ or GFP$^-$ cells from differentiated human pluripotent stem cells (hPSCs) (Figure 4A). While the hESC-NKX2-5 and hiPSC-NKX2-5 samples on day 7 clustered together within the GFP$^+$ group, at the later time points, the samples clustered by cell line suggesting that gene expression profiles of cardiomyocytes do not change significantly between day 14 and day 21 when the cells are maintained in low insulin (LI-)BPEL.

To investigate the correlation between all samples during independent differentiation experiments, biological replicates were averaged and correlation coefficients were calculated. The correlation plots among GFP$^+$ cells between hESC-NKX2-5 and hiPSC-NKX2-5 at matching time points showed concordance and very high correlation coefficients (0.98–0.99) demonstrating that these cells are very similar. As expected, lower correlation values were detected between undifferentiated PSCs and CMs (0.88–0.91) (Figure 4B).

This was also reflected in the number of differentially expressed genes (DEGs) when the different populations were compared (Figure 4C and 4D). The number of DEGs between each time point decreased similarly in both hESC-NKX2-5 and hiPSC-NKX2-5 during the course of the differentiation. As expected, the number of DEGs was the largest between day 0 and day 7 (1625 in hESC-NKX2-5; 1471 in hiPSC-NKX2-5). From day 0 to day 14 of differentiation, the number of DEGs between the two cell types dropped by ~4-fold, indicating very similar processes occur in the development of CPCs and early cardiomyocytes. GO.BP terms related to heart and muscle development, morphogenesis and differentiation were significantly enriched during differentiation in the GFP$^+$ populations, in particular between day 0 and day 7, with the majority of these terms similar between hESC-NKX2-5 and hiPSC-NKX2-5 (Figure 5).

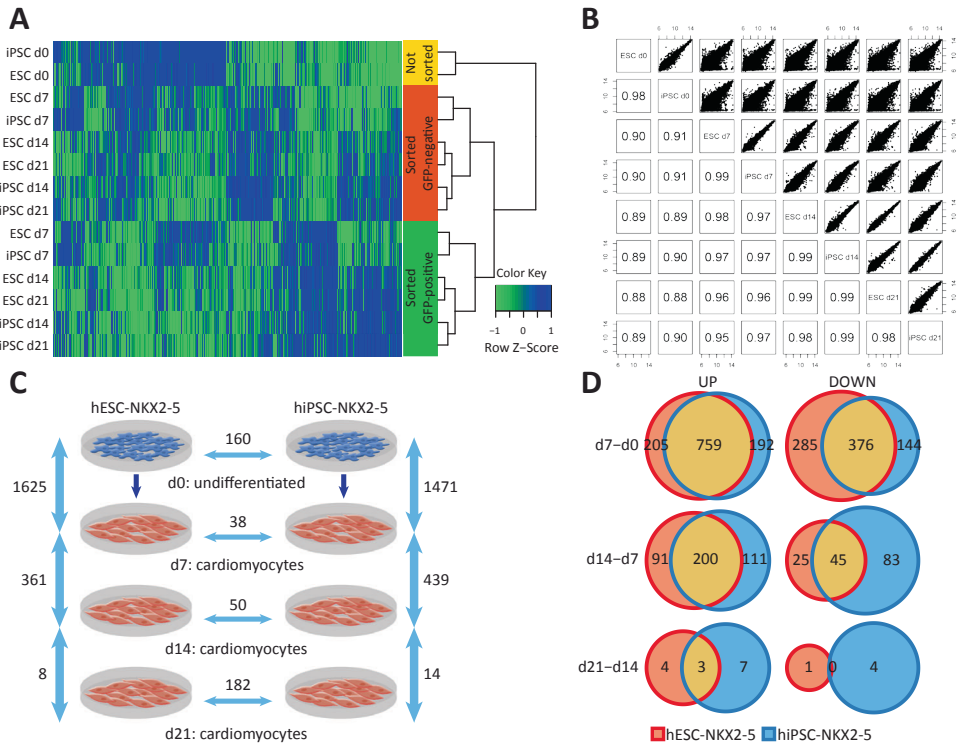

Figure 4. Gene expression analysis of undifferentiated and cardiac hESC- and hiPSC-NKX2-5
A, Three separate clusters are identified by unsupervised hierarchical clustering in hESC- (ESC) and hiPSC-NKX2-5 (iPSC): undifferentiated, sorted GFP$^+$ (NKX2-5$^+$) and sorted GFP$^-$ (NKX2-5$^-$) cells. **B**, Concordance between undifferentiated and GFP-sorted hESC-NKX2-5 (ESC) and hiPSC-NKX2-5 (iPSC) at 4 time points. **C**, Schematic of the number of differentially expressed genes between hESC-NKX2-5 and hiPSC-NKX2-5 at each time point (horizontal) and the number of DEGs between each time point during the differentiation within each cell line (vertical). (n=3) **D**, Venn diagrams demonstrating the number of DEGs during the differentiation, subdivided as up or down regulated genes, and the overlap in transcripts altered between the differentiating NKX2-5-eGFP$^+$ human pluripotent stem cell lines. d indicates day.

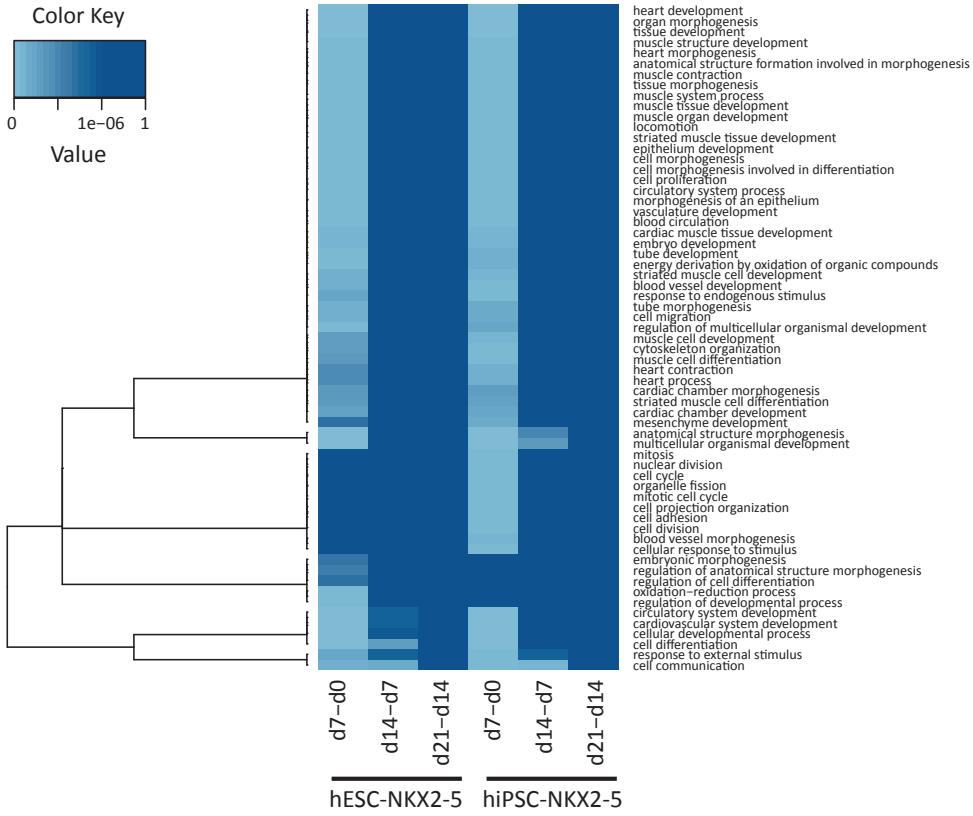

Figure 5. Enriched gene ontology terms belonging to biological processes (GO.BP) for hESC- and hiPSC-NKX2-5

Enriched GO terms of differentially expressed genes during the differentiation of hESC-NKX2-5 and hiPSC-NKX2-5. Significantly enriched GO terms (Fisher's exact test; adjusted *P*-value ≤1e-06) are shown in light blue. d indicates day.

More specifically, Figure 4D shows the number of DEGs during the course of differentiation, subdivided as up or down regulated genes, and the overlap in transcripts altered between the differentiating hPSC lines. The majority of the 759 upregulated genes that overlap between hESC-NKX2-5 and hiPSC-NKX2-5 from day 0 to day 7 were related to developmental processes, metabolism and tissue or muscle morphogenesis; while the 376 genes that were downregulated were enriched for chromatin and DNA organization as well as cell cycle processes (data not shown). As expected, among these overlapping genes, *NKX2-5* is upregulated along with other cardiac-specific genes *MYH6, ACTC1, MYBPC3, TNNC1, MYL7, MYL4, TNNT2 and MYOM1*. Downregulated genes, such as *POU5F1*, were those typically involved in pluripotency or in the cell cycle, (e.g. *CDC25C* and *CDK1*). Nine out of the 10 genes that showed the greatest upregulation in the first 7 days of differentiation were common to

both cell lines (Table 1). Likewise, 7 of the downregulated genes were the same. During the intermediate differentiation stage (day 7 to day 14), genes involved in cardiac development, differentiation and muscle contraction (*MYL2, MYH7, MYH11, VCAM, and CACNA1H*) were significantly upregulated and overlapped between hESC-CMs and hiPSC-CMs, while gastrulation-associated genes and genes involved in the ERK1 and ERK2 signalling cascade (*DUSP5*, *DUSP6* and *FGF8*) were downregulated. Between day 14 and day 21 the number of DEGs in both hESC-CMs and hiPSC-CMs was very low (8 and 14 respectively), indicating the NKX2-5-eGFP$^+$ cell types have a restricted differentiation capacity by day 14.

When hESC-NKX2-5 and hiPSC-NKX2-5 were compared to each other at the various stages of differentiation the number of DEGs was minimal, with between only 38 and 182 transcripts altered out of a total of 16283 transcripts analysed. Between day 14 and day 21 there was a divergence in the NKX2-5-eGFP$^+$ cells derived from either hESC-NKX2-5 or hiPSC-NKX2-5 (Figure 4C) and the largest number of DEGs was detected there. While many of the 182 genes that were significantly differentially expressed between hiPSC-CMs and hESC-CMs at day 21 are involved in muscle development and contraction, no specific gene ontology term or pathway was enriched for. Some of these genes are known to be specific for the heart (*MYOM2, ACTN2*), while others are related to skeletal muscle or muscle in general (*MB, COL11A1*). One group of genes were differentially expressed between the hiPSC and hESC lines at all time points analysed. We postulated that these differences were due to the genetic variability between the two lines rather than differences in cardiac differentiation potential. We therefore identified the DEGs that were only present in the GFP$^+$ population (i.e. absent in the GFP$^-$ and day 0 populations), and compared these between the two cell lines. As in our previous analysis, we did not find any enriched terms at day 7 and day 14, confirming that these populations are very similar. However at day 21 we found the DEGs in the hiPSC-NKX2-5 cells were enriched for biological processes such as muscle development and myoblast differentiation (data not shown), suggesting that these hiPSC-NKX2-5 cells contained a greater proportion of myocytes than the hESC-NKX2-5 cells.

Table 1: Differentially expressed genes at multiple time points for hESC- and hiPSC-NKX2-5

	Downregulated genes				Upregulated genes			
	hESCs		**hiPSCs**		**hESCs**		**hiPSCs**	
Day 0 to 7	*TFF3*	–5,53	*L1TD1*	-4,84	*MYH6*	7,75	*MYH6*	7,87
	POU5F1B	–5,16	*SFRP2*	-4,81	*MYL4*	7,66	*TNNT2*	7,70
	SFRP2	–4,98	*FOXA2*	-4,75	*MYL7*	7,61	*MYL4*	7,67
	FOXA2	–4,95	*POU5F1B*	-4,75	*TNNC1*	7,51	*MYL7*	7,61
	EPHA1	–4,85	*CHGA*	-4,24	*TNNT2*	7,43	*TNNC1*	7,22
	POU5F1	–4,77	*EPHA1*	-4,23	*ACTC1*	7,33	*MYBPC3*	6,74
	L1TD1	–4,52	*T*	-4,09	*MYBPC3*	6,47	*ACTC1*	6,61
	CALCA	–4,29	*HAS3*	-4,01	*MYL3*	6,30	*SMPX*	6,11
	CDH1	–4,07	*POU5F1*	-3,73	*MYOM1*	6,18	*MYOM1*	5,96
	SLC2A1	–3,92	*TFF3*	-3,69	*SMYD1*	6,16	*SMYD1*	5,88
	hESCs		**hiPSCs**		**hESCs**		**hiPSCs**	
Day 7 to 14	*LIN28A*	–3,22	*LIN28A*	-3,19	*LEFTY2*	4,75	*LEFTY2*	4,59
	CBLN2	–2,72	*DGKI*	-2,58	*MT1E*	4,23	*SPHKAP*	4,33
	NPY	–2,62	*PGA5*	-2,29	*MT1A*	4,12	*MYH7*	3,42
	LRRN4	–2,54	*LRRN4*	-2,26	*MT1X*	3,99	*AK4*	3,08
	TNC	–2,39	*GPR19*	-2,20	*SPHKAP*	3,95	*MYL2*	2,96
	APLNR	–2,39	*PGA3*	-2,17	*MYL2*	3,88	*HSPB3*	2,93
	DGKI	–2,37	*FAM19A4*	-2,04	*MYH7*	3,77	*GJA3*	2,64
	HAS2	–2,20	*SLC47A1*	-2,03	*MT2A*	3,53	*TXNIP*	2,52
	PGA5	–2,19	*TNC*	-2,01	*PRSS35*	3,53	*NELL1*	2,51
	PANX2	–2,14	*KEL*	-1,97	*NPPA*	3,31	*NPPA*	2,50
	hESCs		**hiPSCs**		**hESCs**		**hiPSCs**	
Day 14 to 21	*FCN3*	–1,69	*COL22A1*	-1,99	*MYL2*	2,99	*ARPP21*	1,10
			TSPAN32	-1,52	*LGALS3BP*	1,43	*TMEM176B*	1,03
			DYSF	-1,05	*FAM5C*	1,42	*SOSTDC1*	1,30
			COX6A2	-1,01	*TMEM176A*	1,33	*TMEM176A*	1,58
					GOLGA8G	1,12	*ITLN1*	1,51
					TNNI3K	1,05	*FAM5C*	1,33
					SRPK3	1,03	*MYL2*	2,03
							DCN	2,23
							FSTL5	1,45
							H19	2,23

Differentially expressed genes that show the highest significant upregulation between day 0 and 7, day 7 and 14, and day 14 and 21 in both hESC- and hiPSC-NKX2-5.

Discussion

In this study we inserted *eGFP* in the *NKX2-5* gene of hiPSCs to generate a line analogous to an earlier generated hESC-NKX2-5 line [4]. The expression of GFP is thus linked to the cardiac marker *NKX2-5* so that following cardiac differentiation the majority of the GFP^+ cells contract and express endogenous *NKX2-5*. Reporter lines such as these are useful for establishing differentiation protocols that work efficiently over multiple cell lines as they provide a GFP fluorescent readout that can be rapidly evaluated and quantified. Indeed, both the hESC- and hiPSC-NKX2-5 lines were used in the development of the monolayer differentiation protocol described in chapter 4.

By maintaining the undifferentiated PSCs in a chemically defined xeno-free and feeder-free medium, and inducing differentiation using serum-free medium, differentiation procedures were more reproducible and did not require line-to-line optimisation. As previously observed by Chetty et al. [16] we also found that culturing the cells in Essential-8 medium supplemented with 1 % DMSO for 48 hrs prior to starting the differentiation, minimised variability in cardiac differentiation efficiency between independent experiments. Also under these conditions, no modifications were required in the time of addition or the concentration of growth factors and small molecules between the different cell lines. Indeed, when we differentiated both hESCs and hiPSCs using identical defined conditions, the developmental markers as measured by flow cytometry and immunohistochemistry were virtually indistinguishable.

We characterized the hiPSC-NKX2-5 line and compared the GFP^+ cells from both cell lines for cardiomyocyte marker expression [4,6], sarcomere structure, electrophysiology and gene expression. The similar expression of both sarcomeric and cell surface cardiac proteins, indicated that hiPSC-NKX2-5 quantitatively had similar cardiomyogenic potential as hESC-NKX2-5. Both cell lines expressed high percentages of PDGFRa during the mesoderm stage of differentiation and these cells became $PDGFRa^+GFP^+$ from day 7 onwards. While other studies have described the co-expression of KDR and PDGFRa in early cardiovascular cells [17], we only detected low levels of KDR expression and never co-expression with NKX2-5 (GFP), similar to an earlier study using a different hESC-NKX2-5 cell line [5].

When these hPSC-derived cardiomyocytes were compared at the functional level by electrophysiology, they were very similar and showed an immature phenotype. Additionally, the sarcomeres appeared disorganised, unlike the organized alignment of adult cardiomyocytes [18-20]. This immaturity was also reflected in the gene expression profile, with both hESC- and hiPSC-CMs clustering closer to first trimester fetal heart samples than second trimester [21]. Earlier studies have also found immature cardiomyocytes [22,23] and this might be due to the differentiation medium that was used. If these cells were cultured in medium containing thyroid hormone, cardiomyocytes would possibly mature further in culture [21,24].

For this study we chose to use reporter cell lines for their ability to isolate viable NKX2-5-

eGFP$^+$ cells at different stages of differentiation, thereby providing not only a starting point for fate mapping studies and gene profiling experiments, but also the possibility to directly compare such cell populations from both hESC-NKX2-5 and hiPSC-NKX2-5. Here we compared gene expression profiles of enriched populations of hESC-NKX2-5 and hiPSC-NKX2-5 cardiomyocytes at different stages of development. When the PSCs were undifferentiated, less than 1 % of the investigated genes were differentially expressed between hESC-NKX2-5 and hiPSC-NKX2-5.

When comparing the differentiated GFP$^+$ and GFP$^-$ populations from both cell lines, cardiac-related genes were highly upregulated in the GFP$^+$ population. We also found very few DEGs (less than 50) between the GFP$^+$ populations of the hESC-NKX2-5 and hiPSC-NKX2-5 lines at the CPC and early cardiomyocyte stages. However at later stages of differentiation (day 21) the number of DEGs between the two cell lines increased (~180 genes), suggesting divergence in the NKX2-5-eGFP$^+$ population, possibly due to the different genetic backgrounds or epigenetic memory of the somatic cell of origin retained in the hiPSCs. A subset of the DEGs indicated over-representation of genes related to muscle contraction and development in the hiPSCs. Additionally, reduced expression of VCAM1 was observed in the GFP$^+$ hiPSC-NKX2-5 cells at later time points.

Since the first generation of hiPSCs researchers have been investigating how similar these cells are to hESCs, with the majority of the comparisons performed in undifferentiated hPSCs [25]. While hESCs and hiPSCs maintain pluripotency and differentiate through the activation/repression of the same transcription factors and signalling pathways [26], the differentiation capacity of individual cell lines is highly variable, possibly due to epigenetic memory of the cell of origin for hiPSCs, or genetic background [27,28]. Recently, a comparison was made between genetically matched undifferentiated hESCs and hiPSCs. This showed high similarity between the cell lines at both the transcriptional and epigenetic level, suggesting that the genetic background of the cells is a major contributor to the variability observed between different hiPSCs [29]. However, it is possible that epigenetic memory could still contribute to differences in differentiation efficiency to particular cell types. Indeed a study investigating the ability of hiPSC lines with the same genetic background, but derived from different tissue sources (fibroblast- and cardiac progenitor cell-derived), to generate cardiomyocytes, showed differences in their cardiac potential [30]. Since only two lines were compared, the differences observed at later stages of differentiation would be difficult to attribute to either of these possible causes; nevertheless, the low variability in gene expression between the hiPSC and hESC lines at early cardiac differentiation stages is noteworthy, especially since both lines have different genetic backgrounds. This indicates that hiPSCs can be used to study differentiation in early human cardiac development, and are suitable alternatives to hESCs. However, further improvement to functional maturation of both types of hPSC-derived cardiomyocytes is still crucial, and is likely to further improve their accuracy and robustness as models to study human cardiogenesis.

Acknowledgements

We thank D. Ward-van Oostwaard (Anatomy and Embryology, Leiden University Medical Center) for technical assistance and D. de Jong, K. Szuhai, and H. Tanke (Molecular Cell Biology, Leiden University Medical Center) for karyotyping analysis.

Sources of Funding

C.W.v.d.B. was supported by the Netherlands Institute of Regenerative Medicine (NIRM).

References

1 Bilic J., Izpisua Belmonte J.C. Concise review: Induced pluripotent stem cells versus embryonic stem cells: close enough or yet too far apart? *Stem Cells* 30: 33-41 (2012).

2 Narsinh K.H., Sun N., Sanchez-Freire V., Lee A.S. et al. Single cell transcriptional profiling reveals heterogeneity of human induced pluripotent stem cells. *Journal of Clinical Investigation* 121: 1217-1221 (2011).

3 Sepac A., Si-Tayeb K., Sedlic F., Barrett S. et al. Comparison of cardiomyogenic potential among human ESC and iPSC lines. *Cell Transplantation* 21: 2523-2530 (2012).

4 Elliott D.A., Braam S.R., Koutsis K., Ng E.S. et al. NKX2-5(eGFP/w) hESCs for isolation of human cardiac progenitors and cardiomyocytes. *Nature Methods* 8: 1037-1040 (2011).

5 Den Hartogh S.C., Schreurs C., Monshouwer-Kloots J.J., Davis R.P. et al. Dual reporter MESP1(mCherry/w)-NKX2-5(eGFP/w) hESCs enable studying early human cardiac differentiation. *Stem Cells* (2014).

6 Skelton R.J., Costa M., Anderson D.J., Bruveris F. et al. SIRPA, VCAM1 and CD34 identify discrete lineages during early human cardiovascular development. *Stem Cell Research* 13: 172-179 (2014).

7 Yu J., Hu K., Smuga-Otto K., Tian S. et al. Human induced pluripotent stem cells free of vector and transgene sequences. *Science* 324: 797-801 (2009).

8 Davis R.P., Grandela C., Sourris K., Hatzistavrou T. et al. Generation of human embryonic stem cell reporter knock-in lines by homologous recombination. *Current Protocols in Stem Cell Biology* Chapter 5: Unit 5B 1 1 1-34 (2009).

9 Costa M., Dottori M., Sourris K., Jamshidi P. et al. A method for genetic modification of human embryonic stem cells using electroporation. *Nature Protocols* 2: 792-796 (2007).

10 Davis R.P., Costa M., Grandela C., Holland A.M. et al. A protocol for removal of antibiotic resistance cassettes from human embryonic stem cells genetically modified by homologous recombination or transgenesis. *Nature Protocols* 3: 1550-1558 (2008).

11 Szuhai K., Tanke H.J. COBRA: combined binary ratio labeling of nucleic-acid probes for multi-color fluorescence in situ hybridization karyotyping. *Nature Protocols* 1: 264-275 (2006).

12 Costa M., Sourris K., Hatzistavrou T., Elefanty A.G., Stanley E.G. Expansion of human embryonic stem cells in vitro. *Current Protocols in Stem Cell Biology* 5:C: 1C.1.1–1C.1.7 (2008).

13 van den Berg C.W., Elliott D.A., Braam S.R., Mummery C.L., Davis R.P. Differentiation of human pluripotent stem cells to cardiomyocytes under defined conditions. *Methods in Molecular Biology* 1353: 163-180 (2016).

14 Ng E.S., Davis R.P., Stanley E.G., Elefanty A.G. 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* 3: 768-776 (2008).

15 Smyth G.K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology* 3: 1-25 (2004).

16 Chetty S., Pagliuca F.W., Honore C., Kweudjeu A. et al. A simple tool to improve pluripotent stem cell differentiation. *Nature Methods* 10: 553-556 (2013).

17 Kattman S.J., Witty A.D., Gagliardi M., Dubois N.C. 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* 8: 228-240 (2011).

18 Gherghiceanu M., Barad L., Novak A., Reiter I. et al. Cardiomyocytes derived from human embryonic and induced pluripotent stem cells: comparative ultrastructure. *Journal of Cellular and Molecular Medicine* 15: 2539-2551 (2011).

19 Novak A., Barad L., Lorber A., Gherghiceanu M. et al. Functional abnormalities in iPSC-derived cardiomyocytes generated from CPVT1 and CPVT2 patients carrying ryanodine or calsequestrin mutations. *Journal of Cellular and Molecular Medicine* 19: 2006-2018 (2015).

20 Veerman C.C., Kosmidis G., Mummery C.L., Casini S. et al. Immaturity of human stem-cell-derived cardiomyocytes in culture: fatal flaw or soluble problem? *Stem Cells and Development* 24: 1035-1052 (2015).

21 van den Berg C.W., Okawa S., Chuva de Sousa Lopes S.M., van Iperen L. et al. Transcriptome of human foetal heart compared with cardiomyocytes from pluripotent stem cells. *Development* 142: 3231-3238 (2015).

22 Davis R.P., Casini S., van den Berg C.W., Hoekstra M. et al. Cardiomyocytes derived from pluripotent stem cells recapitulate electrophysiological characteristics of an overlap syndrome of cardiac sodium channel disease. *Circulation* 125: 3079-3091 (2012).

23 Bellin M., Casini S., Davis R.P., D'Aniello C. et al. Isogenic human pluripotent stem cell pairs reveal the role of a KCNH2 mutation in long-QT syndrome. *The EMBO Journal* 32: 3161-3175 (2013).

24 Ribeiro M.C., Tertoolen L.G., Guadix J.A., Bellin M. et al. Functional maturation of human pluripotent stem cell derived cardiomyocytes in vitro–Correlation between contraction force and electrophysiology. *Biomaterials* 51: 138-150 (2015).

25 Chin M.H., Mason M.J., Xie W., Volinia S. et al. Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 5: 111-123 (2009).

26 Vallier L., Touboul T., Brown S., Cho C. et al. Signaling pathways controlling pluripotency and early cell fate decisions of human induced pluripotent stem cells. *Stem Cells* 27: 2655-2666 (2009).

27 Kim K., Doi A., Wen B., Ng K. et al. Epigenetic memory in induced pluripotent stem cells. *Nature* 467: 285-290 (2010).

28 Ruiz S., Diep D., Gore A., Panopoulos A.D. et al. Identification of a specific reprogramming-associated epigenetic signature in human induced pluripotent stem cells. *Proceedings of the National Academy of Sciences of the United States of America* 109: 16196-16201 (2012).

29 Choi J., Lee S., Mallard W., Clement K. et al. A comparison of genetically matched cell lines reveals the equivalence of human iPSCs and ESCs. *Nature Biotechnology* 33: 1173-1181 (2015).

30 Sanchez-Freire V., Lee A.S., Hu S., Abilez O.J. et al. Effect of human donor cell source on differentiation and function of cardiac induced pluripotent stem cells. *Journal of the American College of Cardiology* 64: 436-448 (2014).