Fakultät für Medizin

Klinik für Herz- und Gefäßchirurgie des Deutschen Herzzentrums München

Identification of ventricular cardiomyocytes derived from human induced pluripotent stem cells using CRISPR/Cas9

Dr. med. univ. Elda Dzilic

Vollständiger Abdruck der von der
Fakultät für Medizin
der Technischen Universität München zur Erlangung des akademischen Grades
eines Doktors der Medizin (Dr. med.)
genehmigten Dissertation.

Vorsitzender: Prof. Dr. Florian Eyer

Prüfende der Dissertation:

1. apl. Prof. Dr. Markus Krane

2. Prof. Dr. Alessandra Moretti

Die Dissertation wurde am 11.12.2020 bei der Technischen Universität München eingereicht
und durch die Fakultät für Medizin am 13.04.2021 angenommen.

Acknowledgment

Za Babu.

Table of contents

List of abbreviations

ACTB	Beta-Actin
ADP	Action potential duration
AP	Action potential
ATP	Adenosine triphosphate
bFGF	Recombinant Human bFGF-Basic
BLAST	Basic Local Alignment Search Tool
BMP	Bone morphogenetic protein
BSA	Bovine Serum Albumin
Cas9	CRISPR-associated protein 9
cDNA	Complementary deoxyribonucleic acid
C-KIT	Tyrosine-protein kinase KIT
CM	Cardiomyocyte
CRISPR	Clustered Regularly Spaced Inter Palindromic Repeats
CTNT	Troponin-T
CVD	Cardiovascular disease
DAPI	4’,6-diamidino-2-phenylindole
DMEM	Dulbecco’s Modified Eagle’s Medium
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
D-PBS	Dulbecco's Phosphate-Buffered Saline
DSB	Double-strand breaks
DTT	Dithiothreitol
E8	Essential 8

EB	Embryoid body
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
FACS	Fluorescence-activated cell sorting
FBS	Fetal Bovine Serum
FRT	Flippase recognition target
Fwd	Forward
GFP	Green fluorescent protein
HCN4	Hyperpolarization activated cyclic nucleotide gated potassium channel 4
HDR	Homology-directed repair
HEK	Human embryonic kidney
ICC	Immunocytochemistry
iPSC	Induced pluripotent stem cell
IRES	Internal ribosome entry site
IRX4	Iroquois-class homeodomain protein
ISL-1	Insulin gene enhancer protein ISL-1
KLF4	Kruppel-like factor 4
LB	Lysogeny broth
LHA	Left homology arm
LoxP	Locus of X-over P1
LP-FACS	Large particle FACS
MDP	Maximum diastolic potential
MLC2A	Myosin light chain 2
MLC2V	Ventricular myosin light chain 2

MYL2	Myosin light chain 2
NCBI	National Center for Biotechnology Information
Neo	Neomycin
NHEJ	Non-homologous end joining
NKX2.5	Homeobox protein NKX2.5
NPPA	Atrial natriuretic peptide
OCT3/4	Octamer-binding transcription factor 3/4
PCR	Polymerase chain reaction
PGK	Phosphoglycerate kinase 1 promoter
Puro	Puromycin
QRT-PCR	Quantitative real-time polymerase chain reaction
RA	Retinoic acid
Rev	Reverse
RHA	Right homology arm
RNA	Ribonucleic acid
RT	Room temperature
SCA-1	Stem cells antigen-1
SEM	Standard error of the mean
sgRNA	single guide ribonucleic acid
SLN	Sarcolipin
SOC	Super Optimal broth with Catabolite repression
SOX2	sex determining region Y box 2
TALENs	Transcription activator-like effector nucleases
TAZ	Tafazzin
TBX18	T-box transcription factor 18

TBX5	T-box transcription factor 5
TGFB1	TGFB1 Recombinant Human Protein
UTR	Untranslated region
WHO	World Health Organization
Wnt	Wingless and Int-1
ZFNs	Zinc-finger nucleases

Abstract

Differentiation of human induced pluripotent stem cells into cardiomyocytes gives rise to a pool of ventricular, atrial and nodal subtypes of CMs. Different research applications, like regenerative medicine, require the need for a simple and reliable strategy to purify these subtypes.

Genome editing with the new technology CRISPR/Cas9 allows the efficient and cost-effective development of reporter lines. Targeting MYL-2 for the ventricular subtype in human induced pluripotent stem cells enables the identification of ventricular cardiomyocytes. By introducing a fluorescent marker (tdTomato) which is connected to the expression of the MYL-2, the targeted cells are easily identified after differentiation into cardiomyocytes.

Quantitative real-time polymerase chain reaction during cardiomyocyte differentiation shows an increase of cardiac markers and a decrease of pluripotency markers. Simultaneously, an increase of tdTomato expression over the time course of differentiation can be observed. Immunostaining confirms the overlapping expression of MYL-2 and tdTomato. Functional analysis of the reporter line through patch clamp and traction force microscopy display clear characteristics of ventricular cardiomyocytes.

Taken together, the strategy of using CRISPR/Cas9 technology for specific reporter line generation can be a simple approach for cardiomyocyte subtype identification, which will open up new paths for cardiac related research.

1. Introduction

The latest update from the World Health Organization (WHO) confirms again that cardiovascular diseases (CVDs) remain the leading cause of death worldwide. In 2016, 17.9 million people died from CVDs. In other words 31% of all global deaths are due to CVD (World Heart Organization 2017). Without curious minds and meticulous research over the years this number would be even higher. The pioneering, at that time scandalizing work of William Harvey in 1628 describing the cardiovascular system opened the gates for the understanding of heart diseases (Harvey 1628). Since then, research about cardiac diseases led to novel therapies improving outcome and quality of life of affected patients. Nevertheless, an almost uncountable number of questions remains unanswered and waits for eager physicians and researchers to be solved.

1.1 Induced pluripotent stem cells

With the establishment of human induced pluripotent stem cells (iPSCs) by Takahashi *et al.* in 2007 an exciting and highly promising technique was added to the toolbox of cardiovascular research (Takahashi, Tanabe et al. 2007). The investigators were able to reprogram human somatic cells into pluripotent stem cells by adding the four transcription factors *OCT3/4*, *SOX2*, *C-MYC*, and *KLF4*. These iPSCs are very similar to human embryonic stem cells regarding morphology, proliferation, and gene expression. Based on their pluripotency, they are able to differentiate into cells of all three germ layers (Figure 1A) (Takahashi, Tanabe et al. 2007, Moretti, Bellin et al. 2010). This method finally overcame the ethical concerns regarding human embryonic stem cells. Moreover, they hold the possibility to produce patient-specific cells setting a breakthrough in terms of regenerative medicine as well as personalized drug testing and the investigation of underlying patient-specific diseases.

The implementation of this method spread like wildfire in the scientific world. Carvajal-Vergara *et al.*, for example, generated human iPSCs from patients suffering from LEOPARD syndrome (**L**entigines, **E**lectrocardiographic abnormalities, **O**cular hypertelorism, **P**ulmonary valve stenosis, **A**bnormal genitalia, **R**etardation of growth and **D**eafness) (Carvajal-Vergara, Sevilla et al. 2010). The LEOPARD syndrome is a congenital disease with, among others, malformations of the heart and skin. A major disease phenotype in these patients is

hypertrophic cardiomyopathy. The researchers differentiated patient-specific iPSCs into cardiomyocytes and identified mechanisms in the signal pathways of cardiomyocytes which are known to regulate cardiac hypertrophy, and thus were able to reproduce the patient´s phenotype *in vitro*. Others followed, investigating different cardiac diseases like Long-QT or Timothy syndrome (Moretti, Bellin et al. 2010, Yazawa, Hsueh et al. 2011). They were able to successfully mimic the disease phenotypes *in vitro* using patient-specific iPSCs (Figure 1B).

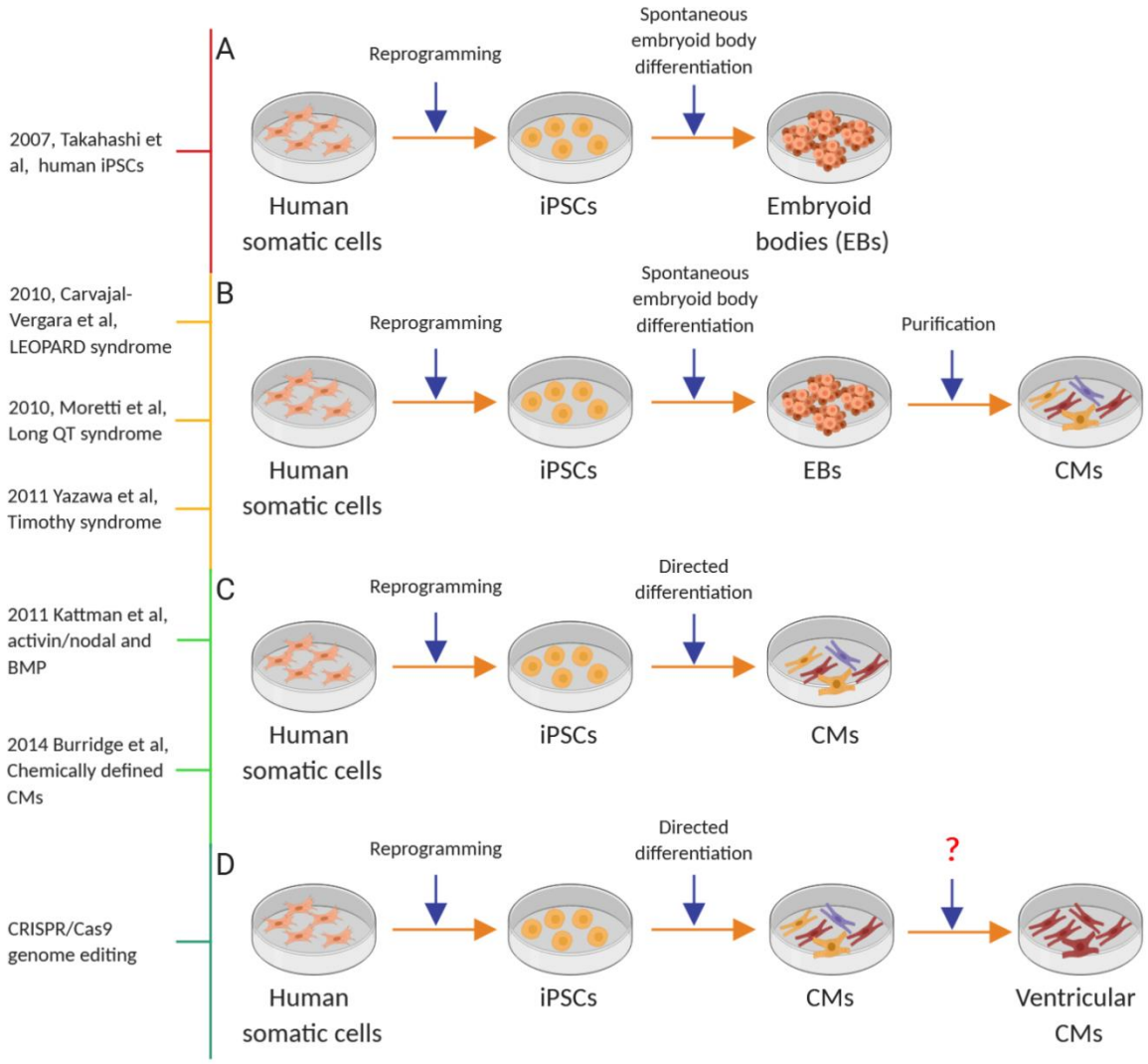


Figure 1: Development of cardiac iPSC research. A) Spontaneous EB differentiation from human iPSCs based on the pioneering work from Takahashi *et al* (Takahashi, Tanabe et al. 2007). **B)** Disease modelling with spontaneous EB differentiation and further purification of cardiomyocytes (CMs). **C)** New methods for directed cardiac differentiation. **D)** Directed cardiac differentiation with the open question of further subtype purification.

However, in order to implement this technique in the clinical setting of regenerative medicine several hurdles had and have to be overcome. For induction of pluripotency mostly

retro- or lentiviral vectors were used in the past. The disadvantage of these vectors is that they incorporate at random into the genome of the target cell and by this possibly destroy essential genes of the reprogrammed cell. Additionally, retroviruses can only successfully infect dividing cells. Viral and non-viral delivery methods that involve only transient expression of genes in a non-integrative manner have been established to circumvent this issue (Okita, Nakagawa et al. 2008, Stadtfeld, Nagaya et al. 2008, Fusaki, Ban et al. 2009, Anokye-Danso, Trivedi et al. 2011). Stadtfeld *et al*. generated iPSCs using non-integrating adenoviruses transiently expressing *Oct4*, *Sox2*, *Klf4*, and *c-Myc* (Stadtfeld, Nagaya et al. 2008). Another approach is to insert the transcription factors via plasmids, which do not integrate into the genome (Okita, Nakagawa et al. 2008). While most of these alternative methods provide a much lower efficiency in producing iPSCs compared to transfection by retroviral vectors, recent introduction of modified RNA and Sendai virus mediated approaches have improved the efficiency of iPSC generation while avoiding the genome integration issue associated with the use of retro- or lentiviruses (Fusaki, Ban et al. 2009, Anokye-Danso, Trivedi et al. 2011). These days the Sendai technique finds widespread application, due to its non-integrative manner, its high efficiency, and easy handling.

Another major problem of iPSC application for regenerative purposes is the risk of tumor formation. After differentiation of iPSCs into the desired cell type, the risk of residual undifferentiated or incompletely differentiated cells remains, enabling uncontrolled cell proliferation. Such an undefined cell population is not suitable for transplantation into patients (Okano and Shiba 2019). One way of preventing tumor formation is to enrich and purify the desired mature cell type. Therefore, methods for purification of these differentiated cell types are required.

1.2 iPSC derived cardiomyocytes

The experience gained from differentiation of human embryonic stem cells were implemented in the work with iPSCs. In the original experiments Takahashi *et al.* performed, inter alia, an embryoid body-mediated differentiation of the iPS cells (Takahashi, Tanabe et al. 2007). This spontaneous *in vitro* differentiation of iPSCs results in a mixture of different cell types, including contracting heart muscle cells, which visually can be identified as beating areas in the cell culture. However, the use of this cell mixture is, as mentioned above, limited due to a high risk of teratoma formation. By manual dissection of the beating

areas or FACS purification by size or markers like C-KIT, SCA-1 or ISL-1, a purity of cardiomyocytes of 50 to 80% can be achieved (Beltrami, Barlucchi et al. 2003, Zwi-Dantsis and Gepstein 2012). Nevertheless, with this level of purity a safe use within the scope of regenerative medicine is not possible. Further limitations of iPSC-derived cardiomyocytes in this field are 1) the necessary cardiomyocyte yield in a clinical setting (Lin, Du et al. 2018), 2) their immunogenicity (Shiba, Gomibuchi et al. 2016, Okano and Shiba 2019) and 3) their immaturity (Scuderi and Butcher 2017, Musunuru, Sheikh et al. 2018). All these obstacles have to be tackled one by one in order to implement hiPSC-derived cardiomyocytes in regenerative medicine. This study focuses on cardiomyocyte subtype identification.

In order to increase cardiomyocyte purity, specific protocols using growth and differentiation factors for directed differentiation into cardiomyocytes were established (Figure 1C). Kattman *et al.* showed that the optimization of levels of Activin/Nodal and BMP leads to efficient cardiac differentiation (Kattman, Witty et al. 2011). Through the optimization of cardiac mesoderm development in mouse and human cultures, it is possible to generate differentiated populations highly enriched for cardiomyocytes (>60%). Despite this progress, little is known about the pathways and macromolecules required for an efficient *in vitro* cardiac differentiation due to the complexity of proprietary media used, resulting in different non-chemically defined protocols (Kattman, Witty et al. 2011, Lian, Hsiao et al. 2012, Lian, Zhang et al. 2013). This consideration led Burridge *et al.* to generate iPS cell derived cardiomyocytes under clear chemically defined conditions on synthetic matrices (Burridge, Matsa et al. 2014). This protocol allows the elucidation of cardiomyocyte macromolecular and metabolic requirements whilst providing a minimally complex system that is based on small molecules controlling the Wnt signaling pathway. The technique results in a cardiomyocytes yield of more than 80% and it is cost-effective as well as efficient (Figure 1C).

All known cardiac differentiation protocols generate a mixture of atrial, ventricular, and nodal cardiomyocytes. The application of a mixed cardiomyocyte population can be problematic in disease modelling and regenerative therapies. For example, when a mixed population of atrial, ventricular, and nodal cardiomyocytes is used to create a ventricular muscle patch, pro-arrhythmic effects may take place due to the incompatibility of the electrical properties of the implanted cells (Shiba, Gomibuchi et al. 2016). On the other hand

the formation of a biological pacemaker would favor an enriched population of nodal cells (Protze, Liu et al. 2017).

Methods to identify the specific subtypes of cardiomyocytes primarily rely on either histological (*e.g.* immunostaining), transcriptional, or single cell patch clamping analyses. Most of these options are cell destructive and do not allow further culture of the identified cells. Thus, it would be of great interest to the cardiac research field to establish new methods that can be employed to identify and purify a specific cardiomyocyte subtype without killing the cells (Figure 1D).

The different cardiomyocyte subtypes not only differ in their electrophysiological and phenotypical properties, but are of course also discriminated by intrinsic genetic properties. Atrial cardiomyocytes, for example, express *NPPA* (atrial natriuretic peptide) or *SLN* (sarcolipin) (Burridge, Matsa et al. 2014). Nodal cardiomyocytes, on the other hand, express *TBX18* (T-box transcription factor 18) and *HCN4* (hyperpolarization activated cyclic nucleotide gated potassium channel 4) (Burridge, Matsa et al. 2014). In recent years, there has been great interest in the use of the atrial and ventricular-specific isoform of members of the myosin family proteins to serve as marker genes for cell identification (Barton and Buckingham 1985). One isoform, the ventricular myosin light chain 2 (*MLC-2v*) is expressed specifically in ventricular cardiomyocytes and is encoded by myosin light chain 2 (*MYL2*) gene located on chromosome 12 (Lee, Ross et al. 1992, Sheikh, Ouyang et al. 2012, Sheikh, Lyon et al. 2015). MLC-2v is an important sarcomeric protein involved in the regulation of calcium-dependent cardiac muscle contraction. It plays an important role in cardiac development and might be involved in the pathogenesis of dilated cardiomyopathy and heart failure (Sheikh, Lyon et al. 2015). Due to its strict expression in ventricular cardiomyocytes during cardiogenesis it can potentially be used as specific marker for ventricular cardiomyocytes.

1.3 Genome editing

Genome editing is the general term for all molecular techniques which directly modify the DNA or rather genes of interest. For this purpose different techniques are available. Besides TALENs (**t**ranscription **a**ctivator-**l**ike **e**ffector **n**ucleases) and Zinc-finger nucleases (ZFNs), the recently developed CRISPR/Cas9 technology promises the most efficient, easy, specific, and

reproducible way to perform genome editing (Jinek, Chylinski et al. 2012, Cong, Ran et al. 2013, Qi, Larson et al. 2013). In 2015, CRSIPR/Cas9 was selected by the *Science* journal as "Breakthrough of the Year" (Frishman 2016). CRISPR stands for Clustered Regularly Spaced Inter Palindromic Repeats and Cas9 simply for the CRISPR-associated protein 9. The CRISPR/Cas9 system originally describes an adaptive immune system in prokaryotes to protect them against foreign DNA (Jinek, Chylinski et al. 2012). The Cas9 enzyme recognizes DNA molecules with foreign genetic information, and splits them at a certain point. Thereby, the foreign DNA is rendered harmless.

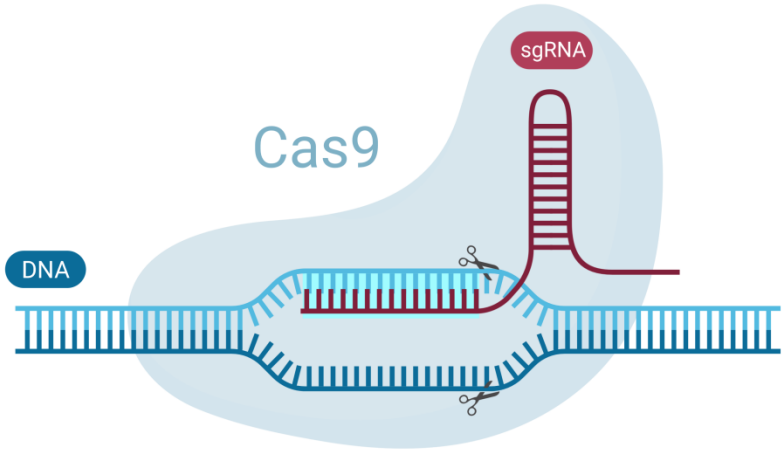


Figure 2: CRISPR/Cas9 technology. The Cas9 is guided through the sgRNA to a specific location within the targeted DNA, where it creates a double strand break.

The CRISPR/Cas9 system has been adapted for inducing sequence-specific double-strand breaks (DSBs) and for targeted genome editing (Jinek, Chylinski et al. 2012). The guide RNA (gRNA) can be custom-designed for nearly any site within the genome at the locus of interest based on sequence homology and the Cas9 protein cuts the DNA at the genome location dictated by the sequence on the gRNA (Figure 2). The generation of DSBs allows the activation of the endogenous cellular repair machinery: non-homologous end joining (NHEJ) or homology-directed repair (HDR). NHEJ is an imprecise mechanism that can result in the generation of nucleotide insertion/deletion (indels) at the targeted site, which may lead to frameshift mutations or premature stop codons, resulting in a dysfunctional protein. On the other hand, HDR is a precise repair mechanism, using the second allele as repair template. This mechanism can be utilized to introduce tailored modifications in the genome by providing an appropriate donor template for repair. A limiting factor of this technique, in the

scope of the cardiovascular field, is that HDR is a mechanism solely occurring in dividing cells. Consequently, this repair mechanism does not take place in cardiomyocytes, as they are post-mitotic cells. Fortunately, first attempts are being made to overcome this hurdle (Waldron 2017).

Nevertheless, CRISPR/Cas9 was adapted in the research of cardiovascular diseases. Wang *et al.* were one of the first groups in the cardiovascular field to utilize CRISPR in order to investigate Barth syndrome, a mitochondrial disorder caused by a mutation of the gene *Tafazzin* (*TAZ)* (Wang, McCain et al. 2014). The investigators introduced a patient-specific *TAZ* mutation in a healthy human iPSC line using CRISPR/Cas9. They differentiated the healthy and edited iPSC line into cardiomyocytes and compared the phenotypes, confirming that the observed abnormalities could only be found in the diseased line. Since then, several other research groups utilized CRISPR/Cas9 for disease modelling, drug screening, and regenerative medicine (Bellin, Casini et al. 2013, Hinson, Chopra et al. 2015, Karakikes, Stillitano et al. 2015, Theodoris, Li et al. 2015, Gupta, Meissner et al. 2016).

2. Objective

The purpose of this study was to create a human iPSC reporter line using CRISPR/Cas9 which allows tracking of ventricular cardiomyocytes. As mentioned above solely ventricular cardiomyocytes express *MLC-2v*. By labelling these cells using CRISPR/Cas9, one can identify the ventricular subtype within the heterogeneous cardiomyocyte population during cardiac differentiation.

The idea is to create a double-strand break right at the end of the *MYL2* gene, to delete the stop codon and to integrate a fluorescent marker that is linked to the gene. By this the fluorescent marker will be transcribed as soon as the gene itself is transcribed. The edited human iPSC line can be differentiated, and ventricular cardiomyocytes can be identified by the fluorescence signal as soon as *MLC-2V* is transcribed.

Establishing this genome-edited human iPSC line is an important step in cardiovascular research. So far, experiments using human iPSC derived cardiomyocytes were done on the pool of different cardiomyocyte subtypes. The generation of the ventricular reporter line allows the optimization of existing differentiation protocols which at best show a mixed cardiomyocyte population of around 85% to 95% comprising ventricular, atrial and nodal cardiomyocytes (Lian, Zhang et al. 2013, Burridge, Matsa et al. 2014). It is known, that these subtypes of cardiomyocytes have different functions, play different functional roles in disease and respond differentially after drug treatment.(Devalla and Passier 2018) Also, regenerative medicine would benefit from the possibility to purify the subtypes, in order to guarantee a specific, direct and safe use in a future clinical setting.

3. Material and methods

3.1 Material

3.1.1 Equipment

Table 1 lists all the equipment used.

Table 1: Equipment

Equipment	Supplier
8-well Chamber Slide with removable wells, #154534	ThermoFisher Scientific, Waltham, MA
ART™ Barrier Reload System, 10 µL, #2139-RTPK	ThermoFisher Scientific, Waltham, MA
ART™ Barrier Reload System, 20 µL, #2149P-RTPK	ThermoFisher Scientific, Waltham, MA
ART™ Barrier Reload System, 200 µL, #2069-RTPK	ThermoFisher Scientific, Waltham, MA
ART™ Barrier Pipette Tips, 1000 µL LR RT, #2179-05-RTPK	ThermoFisher Scientific, Waltham, MA
Corning® CellBIND® Multiple Well Plate size 6 wells, clear bottom flat, sterile, lid, #CLS3335	Merck KGaA, Darmstadt, Germany
Corning® CellBIND® Multiple Well Plate size 12 wells, clear bottom flat, sterile, lid, #CLS3336	Merck KGaA, Darmstadt, Germany
Corning® CellBIND® Multiple Well Plate size 24 wells, clear bottom flat, sterile, lid, #CLS3337	Merck KGaA, Darmstadt, Germany
Corning™ Falcon™ Test Tube with Cell Strainer Snap Cap, #10585801	Fisher Scientific, Loughborough, UK
Eppendorf Safe-Lock Tubes, 1.5 mL, Eppendorf Quality™, #0030120086	Eppendorf AG, Hamburg, Germany
Eppendorf Safe-Lock Tubes, 2 mL, Eppendorf Quality™, #0030120094	Eppendorf AG, Hamburg, Germany
Falcon® 15 mL Polystyrene Centrifuge Tube, Conical Bottom, with Dome Seal Screw Cap, #352095	Corning Incorporated, Corning, NY
Falcon® 50 mL High Clarity PP Centrifuge Tube, Conical Bottom, Sterile, #352070	Corning Incorporated, Corning, NY
Falcon™ Disposable Polystyrene Serological Pipets, Sterile, 5 mL, #10468312	Fisher Scientific, Loughborough, UK
Falcon™ Disposable Polystyrene Serological Pipets, Sterile, 10 mL,#10282371	Fisher Scientific, Loughborough, UK
Falcon™ Disposable Polystyrene Serological Pipets, Sterile, 25 mL, #10683282	Fisher Scientific, Loughborough, UK
Falcon™ Disposable Polystyrene Serological Pipets, Sterile, 50 mL, #10282441	Fisher Scientific, Loughborough, UK
Falcon® Round-Bottom Tubes with Cell Strainer Cap, 5 mL, #10585801	Fisher Scientific, Loughborough, UK
Greiner CELLSTAR® 384 well plates, #M3937	Merck KGaA, Darmstadt, Germany
Nalgene™ CryoVial for long-term storage, 2 mL, #5000-0020	ThermoFisher Scientific, Waltham, MA
MicroAmp™ Optical 96-Well Reaction Plate,	ThermoFisher Scientific, Waltham, MA

#N8010560	
MicroAmp™ Optical Adhesive Film, #4311971	ThermoFisher Scientific, Waltham, MA
PARAFILM® M, #P7793-1EA	Merck KGaA, Darmstadt, Germany
PCR Strips 0.2 mL 8-Tube and Domed Cap Strips, high profile, clear #TBC0802	Bio Rad, Hercules, CA
Petri dishes, polystyrene (100 mm x 15 mm), #P5856-500EA	Merck KGaA, Darmstadt, Germany
Stericup-GV Sterile Vacuum Filtration System, # SCGVU02RE	Merck KGaA, Darmstadt, Germany
Thermo Scientific™ Richard-Allan Scientific™ Cover Glas, #22-050-218	ThermoFisher Scientific, Waltham, MA
VWR® Soft Nitrile Examination Gloves, #89038-270	VWR International, Radnor, PA

3.1.2 Reagents

Table 2 lists all the reagents used in this study.

Table 2: Reagents

Reagent	**Supplier**
1 Kb Plus DNA Ladder, #10787018	ThermoFisher Scientific, Waltham, MA
4-Aminopyridine, #A78403	Merck KGaA, Darmstadt, Germany
Acetic acid, CH_3CO_2H, #A6283	Merck KGaA, Darmstadt, Germany
Agarose, #A4718-25G	Merck KGaA, Darmstadt, Germany
Ampicillin sodium salt, #A0166-5G	Merck KGaA, Darmstadt, Germany
ATP Solution (100 mM), #R0441	ThermoFisher Scientific, Waltham, MA
B-27™ Plus Supplement (50X), #A3582801	ThermoFisher Scientific, Waltham, MA
B-27™ Supplement, minus insulin (50X), #A1895601	ThermoFisher Scientific, Waltham, MA
Bambanker ® Freezing Medium, #302-14681	FUJIFILM Wako Pure Chemical Corporation, Osaka, Japan
Blebbistatin, #B0560	Merck KGaA, Darmstadt, Germany
Bovine Serum Albumin, #15561020	ThermoFisher Scientific, Waltham, MA
CHIR99021, #72052	STEMCELL Technologies Inc., Grenoble, France
Corning Matrigel Matrix, #354277	Corning Incorporated, Corning, NY
D-(+)-Glucose solution, #G8644	Merck KGaA, Darmstadt, Germany
Dimethylsulfoxide (DMSO), #10127403	ThermoFisher Scientific, Waltham, MA
Dispase (1 U/mL), #07923	STEMCELL Technologies Inc., Grenoble, France
DNA Gel Loading Dye (6X), #R0611	ThermoFisher Scientific, Waltham, MA
dNTP Mix (10 mM each), #R0191	ThermoFisher Scientific, Waltham, MA
Dulbecco's Phosphate-Buffered Saline without calcium chloride, without magnesium chloride (D-PBS), #A1285601	ThermoFisher Scientific, Waltham, MA
DTT, 1 M, #P2325	ThermoFisher Scientific, Waltham, MA
Ethylenediaminetetraacetic acid, EDTA, #E6758-100G	Merck KGaA, Darmstadt, Germany
Ethanol 96%, CH_3CH_2OH, #16368	Merck KGaA, Darmstadt, Germany
Ethidium bromide solution, #E1510	Merck KGaA, Darmstadt, Germany
Fetal Bovine Serum (FBS), charcoal stripped, #12676029	ThermoFisher Scientific, Waltham, MA

FluoroBrite™ DMEM, #A1896701	ThermoFisher Scientific, Waltham, MA
FluoVolt™ Membrane Potential, #F10488	ThermoFisher Scientific, Waltham, MA
G 418 disulfate salt solution, #G8168	Merck KGaA, Darmstadt, Germany
Gibco Recombinant Human bFGF-Basic (bFGF), #PHG0367	ThermoFisher Scientific, Waltham, MA
GlutaMAX™ Supplement (100x), #35050061	ThermoFisher Scientific, Waltham, MA
Glycerin, #56-81-5	Merck KGaA, Darmstadt, Germany
Heparin sodium salt from porcine intestinal mucosa, #H3393-250KU	Merck KGaA, Darmstadt, Germany
Hoechst 33258 solution, #94403	Merck KGaA, Darmstadt, Germany
Human Insulin Powder, #800-112P	GeminiBio, West Sacramento, CA
Hydrochloric acid solution, #H3162	Merck KGaA, Darmstadt, Germany
L-Ascorbic acid, #A92902-25G	Merck KGaA, Darmstadt, Germany
L-Glutamine (200 mM), #25030081	ThermoFisher Scientific, Waltham, MA
LB Broth with agar, #L3147	Merck KGaA, Darmstadt, Germany
Methanol anhydrous, 99.8%, CH_3OH, #322415	Merck KGaA, Darmstadt, Germany
Mounting Medium with DAPI, #ab104139	Abcam, Cambridge, UK
Nuclease-Free Water, #AM9920	ThermoFisher Scientific, Waltham, MA
Paraformaldehyde, reagent grade, crystalline. #P6148	Merck KGaA, Darmstadt, Germany
Plasmid-Safe ATP-Dependent DNase, #E3101K	Lucigen, Hilden, Germany
Pluronic® F-127, #P2443	Merck KGaA, Darmstadt, Germany
Puromycin dihydrochloride, #P8833	Merck KGaA, Darmstadt, Germany
Restriction enzyme BpiI (BbsI) (10 U/µL), #ER1011	ThermoFisher Scientific, Waltham, MA
RNase-Free DNase Set, #79254	Qiagen, Hilden, Germany
Rock-Inhibitor Y-27632, #72302	STEMCELL Technologies Inc., Grenoble, France
SOC Outgrowth Medium, #B9020S	New England Biolabs, Ipswich, MA
Sodium hydroxide solution (10.0 N), #SX0607N	Merck KGaA, Darmstadt, Germany
Sodium phosphate dibasic, Na_2HPO_4, #S3264	Merck KGaA, Darmstadt, Germany
Sodium selenite, #214485	Merck KGaA, Darmstadt, Germany
StemPro™ Accutase™ Cell Dissociation Reagent, #A1110501	ThermoFisher Scientific, Waltham, MA
T4 DNA Ligase Reaction Buffer, #B0202S	New England Biolabs, Ipswich, MA
T4 Polynucleotide Kinase, #M0201S	New England Biolabs, Ipswich, MA
T7 DNA Ligase, #M0318S	New England Biolabs, Ipswich, MA
Tango Buffer (10X), #BY5	ThermoFisher Scientific, Waltham, MA
TGFB1 Recombinant Human Protein (TGFB1), #PHG9211	ThermoFisher Scientific, Waltham, MA
Thiazovivin, #SML1045	ThermoFisher Scientific, Waltham, MA
Transferrin human, #T3309	ThermoFisher Scientific, Waltham, MA
Tris Acetate-EDTA buffer, #T9650	Merck KGaA, Darmstadt, Germany
Tris-Glycine-SDS Buffer, #T7777-1L	Merck KGaA, Darmstadt, Germany
Triton™ X-100, #X100	Merck KGaA, Darmstadt, Germany
Trypan Blue solution, #T8154	Merck KGaA, Darmstadt, Germany
TrypLE™ Select Enzyme (1X), no phenol red, #12563011	ThermoFisher Scientific, Waltham, MA
Wnt-C59, Wnt Antagonist, #ab142216	Abcam, Cambridge, UK

3.1.3 Purchased kits

Table 3 lists all the purchased kits used. The reagents were stored and used as recommended by the suppliers.

Table 3: Kits

Kit	Supplier
DNeasy Blood & Tissue Kit, #69504	Qiagen, Hilden, Germany
GoTaq® DNA Polymerase, #M3001	Promega Corporation, Fitchburg, WI
High-Capacity cDNA Reverse Transcription Kit, #4368814	ThermoFisher Scientific, Waltham, MA
Human Stem Cell Nucleofector® Kit1, #VPH-5012	Lonza Cologne GmbH, Cologne, Germany
HotStar HiFidelity Polymerase Kit, #202602	Qiagen, Hilden, Germany
Power SYBR™ Green PCR Master Mix, #4368577	ThermoFisher Scientific, Waltham, MA
QIAGEN Plasmid Mini Kit, #12123	Qiagen, Hilden, Germany
QIAGEN Plasmid Midi Kit, #12143	Qiagen, Hilden, Germany
QIAquick PCR Purification Kit, #28104	Qiagen, Hilden, Germany
RNeasy Mini Kit, #74104	Qiagen, Hilden, Germany

3.1.4 Antibodies

The antibodies used in this study were used and stored as recommended by the suppliers. Tables 4 and 5 list all the primary and secondary antibodies used, respectively. All primary antibodies were used at a 1:200 dilution. The secondary antibodies were used at a 1:500 dilution. Dilutions were done with D-PBS containing 3% bovine serum albumin (BSA).

Table 4: Primary antibodies

Primary antibody	Supplier
Anti-Nanog antibody [EPR2027(2)], #ab109250	Abcam, Cambridge, UK
Anti-Human TRA-1-81 Antibody, Clone TRA-1-81, #60065	STEMCELL Technologies Inc., Grenoble, France
Anti-Cardiac Troponin I antibody, #ab47003	Abcam, Cambridge, UK
Anti-Myosin Light Chain 2 antibody, #ab79935	Abcam, Cambridge, UK
Anti-F-actin antibody [NH3], #ab205	Abcam, Cambridge, UK

Table 5: Secondary antibodies

Secondary antibody	Supplier
Goat Anti-Mouse IgG H&L (Alexa Fluor® 594), #ab150120	Abcam, Cambridge, UK
Goat Anti-Rabbit IgG H&L (Alexa Fluor® 488), #ab150077	Abcam, Cambridge, UK
Donkey Anti-Rabbit IgG H&L (Alexa Fluor® 647), #ab150075	Abcam, Cambridge, UK

3.1.5 Oligonucleotides

Oligonucleotide design for polymerase chain reaction (PCR), quantitative real-time polymerase chain reaction (qRT-PCR) and single-guide RNAs (sgRNAs) was performed using the website benchling.com (https://www.benchling.com/). The software determined the specificity of the primers and made sure that the designed qRT-PCR primers detected all the existing variants. The off-target locations were also determined using benchling.com based on the work from Lin *et al* (Lin, Cradick et al. 2014). The designed oligonucleotides were purchased via the Protein and Nucleic Acid Facility at Stanford School of Medicine and stored as recommended at -20°C. Tables 6 and 7 list all the primers used in this study. Table 8 lists the tested sgRNAs.

Table 6: PCR primers

Target	Forward (5'→3')	Reverse (5'→3')
Donor sequence insertion in *MYL-2* locus	TTTGTGATCCCAGTGGAAGGCT	GTTCATGCGCTTCAAGGTGCGC
WT *MYL-2* expression	TTTGTGATCCCAGTGGAAGGCT	GGTACTCGGGGGAGAGAGAT
Off-Target Locus 1	GGTGAATCAGAAAGGCAGAA	GCGTAAGTGGATTGTGATGAG
Off-Target Locus 2	TTTTATCTTGAAAATGTTTGCCTA	TGCTGTAGCCCAGAATGATG
Off-Target Locus 3	TTCTGATGAAGCTATTCTGCTCA	CCGGTTGAGAATCACTGGTC
Off-Target Locus 4	TTGGCCCTTTAAATTCAATACTT	ATGGTCCCCTCCTCCTTTAC
Off-Target Locus 5	ACAGGGCTTACTCTGCGTTT	CTGGCCCAGAGGTCACTTAG
Off-Target Locus 6	CCATCTCCTGACCTCGTGAT	GCTGCACCTTCCAGAACAT
Off-Target Locus 7	CTGACACAGAGCCCCTCAC	CTGAGGCACAGGAACAGACA
Off-Target Locus 8	AGCAAAATAGCAACTCCCATAA	TGGAGAGTAAGATTCTCAGAAAGGA
Off-Target Locus 9	TCAGTCTTCCCAGAATGACTCTC	TCTCCAAATTAAAATGAGTTTCTGC
Off-Target Locus 10	TCCCAGCTTTGTGATTCTCC	TGCCTAAAGCAGTGGGTCTT
Off-Target Locus 11	GCCTCACCCAAATTGTCATT	GCACCAATGACCCTTTCCTA
Off-Target Locus 12	TCAGCATGCTTCTCCCCTAC	TTTGAGGATGAGGGGTTTGT
Off-Target Locus 13	GAGACCCTCGTGGTGTCAGT	TTTCACCATGTTGGTCAGGA
Off-Target Locus 14	ACTCCTGCCCTCAGGTGAT	GCCCTACAACGAGGATCAGA
Off-Target Locus 15	CACATGAATGAGTGCTGAACA	CCACCAGACACCCTCCACTA
Off-Target Locus 16	CCTCCCTGCTCCACATACAT	TCACACAGCCAGTTCTGAGG
sgRNA cutting efficiency	TTTGTGATCCCAGTGGAAGGCT	GCGCACCTTGAAGCGCATGAAC
U6 primer	GAGGGCCTATTTCCCATGATTCC	

Table 7: qRT-PCR primers

Target	Forward (5'→3')	Reverse (5'→3')
NANOG	TTTGTGGGCCTGAAGAAAACT	AGGGCTGTCCTGAATAAGCAG
OCT-4	CTTGAATCCCGAATGGAAAGGG	GTGTATATCCCAGGGTGATCCTC
MYL-2	TTGGGCGAGTGAACGTGAAAA	CCGAACGTAATCAGCCTTCAG
CTNT	GGAGGAGTCCAAACCAAAGCC	TCAAAGTCCACTCTCTCTCCATC
tdTomato	ACCATCGTGGAACAGTACGAG	CTTGAAGCGCATGAACTCTTT

BETA-ACTIN	CATGTACGTTGCTATCCAGGC	CTCCTTAATGTCACGCACGAT

Table 8: sgRNA targeting the *MYL-2* locus

	Upper strand (5'→3')	Lower strand (5'→3')
sgRNA1	CACCGACGGAGAAGAGAAGGACTAG	AAACCTAGTCCTTCTCTTCTCCGTC
sgRNA2	CACCGCGGAGAAGAGAAGGACTAGG	AAACCCTAGTCCTTCTCTTCTCCGC
sgRNA3	CACCGAGAAGAGAAGGACTAGGAGG	AAACCCTCCTAGTCCTTCTCTTCTC

3.1.6 Plasmids

The following plasmids were used in this study.

Donor construct:

The donor construct is required for the homologous recombination during genome editing using the CRISPR/Cas9 technology. After choosing the correct sequence for the goal of this study using the National Center for Biotechnology Information (NCBI) database, the service company GenScript (New Jersey, USA) cloned the sequence into the commercially available pUC57 vector (Figure 3).

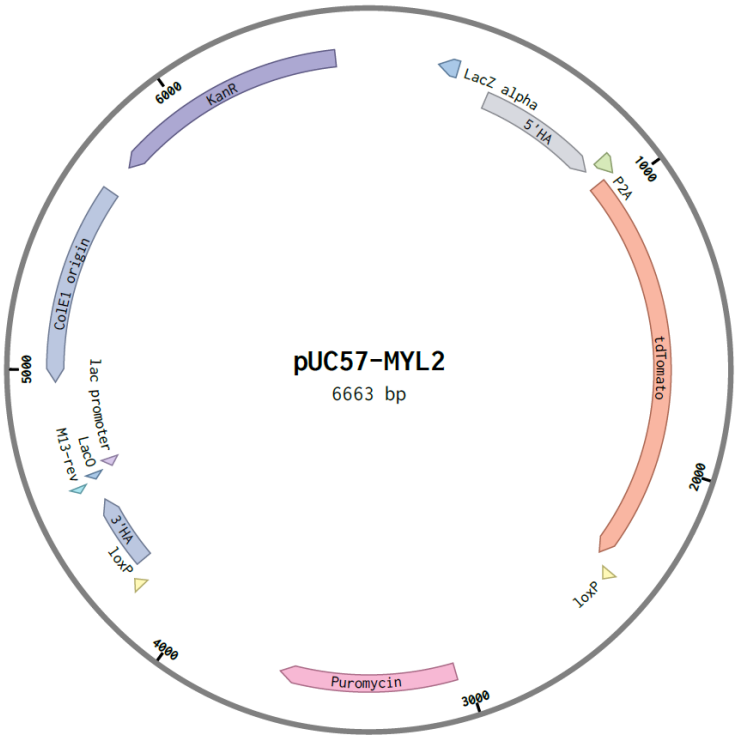


Figure 3: Schematic vector map of pUC57-MYL2 donor vector.

The sequence of the donor construct is as follows: Left-homology arm – **P2A linker** – tdTomato – **loxP** – Phosphoglycerate Kinase (PGK) promoter and Puromycin – **loxP** – right-homology arm. The P2A linker is a self-cleaving peptide that allows the simultaneous expression of *MYL2* and *tdTomato,* while also enabling the cleavage of the two proteins (Wang, Wang et al. 2015). The included loxP sites allow the excision of the flanked sequence within the two loxP sites by adding Cre recombinase (Orban, Chui et al. 1992, Lee and Saito 1998). This allows the removal of the puromycin resistance cassette, if necessary.

GGAAGAAGTTGAGCCAAGATTTAGGCCCAAAGATCGGCACTGCTGACTACCCAGGGAGGGAGCTTATT
CTACAGCTCTCCTGAGAGGTACCAAGTTGGGCGAGATGGCCGTGCCTTTATACCCCCACACCAATCCC
ATCCGCTGGCTGTGGGCTGCCCCAGAAAGGGCATGACCTTGGGCGAGGGGGCCCTGCAGCAGCTGGGG
AAGTCCTTCCCCGGAGGACCCTCTGAGTGGCCCATGTTCCCATCCCCCACCCCCCAGCTATCCGTCCA
CTCAGGCCCCGCCCACAGCCCCCCACCCTCCGTCTCAGTTCCCCTCCCCTCGTCCTTAGCACGTGTTG
CTGGCTCATTGCAGGTTGACCAGATGTTCGCCGCCTTCCCCCCTGACGTGACTGGCAACTTGGACTAC
AAGAACCTGGTGCACATCATCACCCACGGAGAAGAGAAGGAC**GCTAGCGCCACTAACTTCTCCCTGTT**
GAAACAAGCAGGGGATGTCGAAGAGAATCCCGGGCCAATGGTGAGCAAGGGCGAGGAGGTCATCAAAG
AGTTCATGCGCTTCAAGGTGCGCATGGAGGGCTCCATGAACGGCCACGAGTTCGAGATCGAGGGCGAG
GGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGCGGCCCCCTGCC
CTTCGCCTGGGACATCCTGTCCCCCCAGTTCATGTACGGCTCCAAGGCGTACGTGAAGCACCCCGCCG
ACATCCCCGATTACAAGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAG
GACGGCGGTCTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCACGCTGATCTACAAGGTGAA
GATGCGCGGCACCAACTTCCCCCCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCT
CCACCGAGCGCCTGTACCCCCGCGACGGCGTGCTGAAGGGCGAGATCCACCAGGCCCTGAAGCTGAAG
GACGGCGGCCACTACCTGGTGGAGTTCAAGACCATCTACATGGCCAAGAAGCCCGTGCAACTGCCCGG
CTACTACTACGTGGACACCAAGCTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGT
ACGAGCGCTCCGAGGGCCGCCACCACCTGTTCCTGGGGCATGGCACCGGCAGCACCGGCAGCGGCAGC
TCCGGCACCGCCTCCTCCGAGGACAACAACATGGCCGTCATCAAAGAGTTCATGCGCTTCAAGGTGCG
CATGGAGGGCTCCATGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGG
GCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGCGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCC
CCCCAGTTCATGTACGGCTCCAAGGCGTACGTGAAGCACCCCGCCGACATCCCCGATTACAAGAAGCT
GTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGTCTGGTGACCGTGA
CCCAGGACTCCTCCCTGCAGGACGGCACGCTGATCTACAAGGTGAAGATGCGCGGCACCAACTTCCCC
CCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCACCGAGCGCCTGTACCCCCG
CGACGGCGTGCTGAAGGGCGAGATCCACCAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGG
AGTTCAAGACCATCTACATGGCCAAGAAGCCCGTGCAACTGCCCGGCTACTACTACGTGGACACCAAG
CTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAGCGCTCCGAGGGCCGCCA
CCACCTGTTCCTGTACGGCATGGACGAGCTGTACAAGTAAGAATTCCGATCATATTCAATAACCCTTA
AT**ATAACTTCGTATAATGTATGCTATACGAAGTTAT**TAGGTCTGAAGAGGAGTTTACGTCCAGCCAAG
CTTAGGATCTCGACCTCGAAATTCTACCGGGTAGGGGAGGCGCTTTTCCCAAGGCAGTCTGGAGCATG
CGCTTTAGCAGCCCCGCTGGCACTTGGCGCTACACAAGTGGCCTCTGGCCTCGCACACATTCCACATC
CACCGGTAGCGCCAACCGGCTCCGTTCTTTGGTGGCCCCTTCGCGCCACCTTCTACTCCTCCCCTAGT
CAGGAAGTTCCCCCCCGCCCCGCAGCTCGCGTCGTGCAGGACGTGACAAATGGAAGTAGCACGTCTCA
CTAGTCTCGTGCAGATGGACAGCACCGCTGAGCAATGGAAGCGGGTAGGCCTTTGGGGCAGCGGCCAA
TAGCAGCTTTGCTCCTTCGCTTTCTGGGCTCAGCAGCTGGGAAGGGTGGGTCCGGGGGCGGGCTCAGG
GGCGGGCTCAGGGGCGGGGCGGGCGCCCGAAGGTCCTCCGGAGGCCCGGCATTCTGCACGCTTCAAAA
GCGCACGTCTGCCGCGCTGTTCTCCTCTTCCTCATCTCCGGGCCTTTCGACCTGCATCCATCTAGATC
TCGAGCAGCTGAAGCTTACCATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTC
CCCAGGGCCGTACGCACCCTCGCCGCCGCGTTCGCCGACTACCCCGCCACGCGCCACACCGTCGATCC
GGACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCG
GCAAGGTGTGGGTCGCGGACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGCG
GGGGCGGTGTTCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCA

ACAGATGGAAGGCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGCGTGGTTCCTGGCCACCGTCGGCG
TCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGAG
CGCGCCGGGGTGCCCGCCTTCCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTCTACGAGCGGCTCGG
CTTCACCGTCACCGCCGACGTCGAGGTGCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCG
GTGCCTGACGCCCGCCCCACGACCCGCAGCGCCCGACCGAAAGGAGCGCACGACCCCATGCATCGATG
ATATCAGATCCCCGGGATGCAGAAATTGATGATCTATTAAACAATAAAGATGTCCACTAAAATGGAAG
TTTTTCCTGTCATACTTTGTTAAGAAGGGTGAGAACAGAGTACCTACATTTTGAATGGAAGGATTGGA
GCTACGGGGGTGGGGGTGGGGTGGGATTAGATAAATGCCTGCTCTTTACTGAAGGCTCTTTACTATTG
CTTTATGATAATGTTTCATAGTTGGATATCATAATTTAAACAAGCAAAACCAAATTAAGGGCCAGCTC
ATTCCTCCCACTCATGATCTATAGATCTATAGATCTCTCGTGGGATCATTGTTTTTCTCTTGATTCCC
ACTTTGTGGTTCTAAGTACTGTGGTTTCCAAATGTGTCAGTTTCATAGCCTGAAGAACGAGATCAGCA
GCCTCTGTTCCACATACACTTCATTCTCAGTATTGTTTTGCCAAGTTCTAATTCCATCAGAAGCTGGT
CGAGATCCGGAACCCTTAAT**ATAACTTCGTATAATGTATGCTATACGAAGTTAT**TAGGTCCCTCGAAG
AGGTTCACTAGGCGCGCCGAGGGGGCTCGCTGCTGCGCCCTGGGCTCGTCTTTGCAGAGTGGTCCCTG
CCCTCATCTCTCTCCCCCGAGTACCGCCTCTGTCCCTACCTTGTCTGTTAGCCATGTGGCTGCCCCAT
TTATCCACCTCCATCTTCTTTGCAGCCTGGGTGGCTATGGGTACTTCGTGGCCGCACATCCTACAGTT
GGAAATCCATCCAGAGGCCATGTTCCAATAAACAGGAGGTCGTGTATTTGGTCACGACATTTCTCTGA
CAAA

pX458 Plasmid

The pX458 plasmid (pSpCas9(BB)-2A-GFP) was a gift from Feng Zhang (Addgene plasmid #48138 ; http://n2t.net/addgene:48138 ; RRID:Addgene_48138). It contains the sequence for the Cas9 enzyme combined with the GFP fluorophore and a sgRNA cloning site (Ran, Hsu et al. 2013).

Figure 4: Schematic vector map of pX458.

3.1.7 Media and buffers

The media and buffers required for the experiments were either bought commercially or prepared at the bench. Table 9 lists all commercial media. The described solutions mixed at the bench were all filtered using Stericup-GV Sterile Vacuum Filtration System.

Table 9: Media

Medium	Supplier
DMEM, high glucose, #11965092	ThermoFisher Scientific, Waltham, MA
DMEM/F-12, #11320033	ThermoFisher Scientific, Waltham, MA
DMEM/F-12, GlutaMAX™ Supplement, #10565018	ThermoFisher Scientific, Waltham, MA
RPMI 1640 Medium, no glucose, #11879020	ThermoFisher Scientific, Waltham, MA
RPMI Medium 1640 - L-Glutamine, #11875-085	ThermoFisher Scientific, Waltham, MA
RPMI 1640 Medium, GlutaMAX™ Supplement, #61870036	ThermoFisher Scientific, Waltham, MA

Ampicillin sodium salt:
Ampicillin sodium salt was dissolved to a concentration of 100 mg/mL using double-distilled water. The solution was aliquoted and stored at -20°C for up to six months.

B27-Insulin + RPMI1640 + CHIR99021:
B27-Insulin + RPMI1640 + CHIR99021 was made by adding 6 µM of CHIR99021 to one liter of RPMI 1640 + B27-Insulin medium. The solution was stored at 4°C for up to four weeks.

B27-Insulin + RPMI1640 + Wnt-C59:
B27-Insulin + RPMI1640 + Wnt-C59 was made by adding 2 µM of Wnt-C59 to one liter of RPMI 1640 + B27-Insulin medium. The solution was stored at 4°C for up to four weeks.

Bovine Serum Albumin (BSA) solution:
BSA was dissolved with 1x D-PBS to a final concentration of 3% or 4%. The solution was stored at 4°C for up to four weeks.

CHIR99021:
CHIR99021 was dissolved in DMSO to a final concentration of 6 mM, filter-sterilized, aliquoted and stored at -80°C for up to six months.

Dulbecco's phosphate-buffered saline (D-PBS):
10x D-PBS was diluted using double-distilled water and autoclaved. The solution was stored at room temperature (RT) for up to four weeks.

Ethanol:

Ethanol (96%) was diluted to a final concentration of 70% using double-distilled water. The solution was stored at RT.

Essential 8 supplement:

Essential 8 (E8) medium was made at the bench by combining 25 mL nuclease-free water with 3.2 g of L-ascorbic acid in a 50 mL Falcon tube. Besides, 1 g of insulin was added to 23 mL nuclease-free water in another 50 mL Falcon tube. Afterwards, the pH was adjusted slowly to 3 using 1 N hydrochloric acid, and raised again to pH 7.4 by slowly adding 10 N sodium hydroxide. Next, 250 mg transferrin, 500 µL of a 10 mg/mL heparin sodium salt solution, and 500 µL of a 1.4 mg/mL sodium selenite solution were added. This solution was mixed with the 25 mL L-ascorbic acid 2-phosphate solution. In a different 50 mL Falcon tube 5 mg bFGF and 100 µg TGFB1 were dissolved in 25 mL nuclease-free water. The two solutions were mixed, filtered and aliquoted into 1.5 mL portions. These were stored at -20°C for up to 6 months.

Essential 8 medium:

Essential 8 medium was prepared by adding an aliquot of the Essential 8 supplement into one liter of DMEM/F12. The medium was stored at 4°C up to one month.

E8-Y medium:

E8-Y medium was prepared by adding 10 µM of Rock-Inhibitor Y-27632 into 1000 mL of E8 medium. The medium was stored at 4°C up to one month.

Lysogeny broth (LB):

LB was made by adding 20 g LB powder to 1 liter of deionized water. The solution was autoclaved and stored at 4°C for up to four weeks.

Lysogeny broth (LB) agar plate with Ampicillin:

LB agar was made by adding 40 g LB powder to 1 liter of deionized water. The solution was heated in the microwave oven until the powder was dissolved, and autoclaved. One milliliter of ampicillin (100 mg/mL) was added to one liter of the LB agar. Twenty mL of this solution were added to one 100 mm x 15 mm petri dish. The dishes were sealed with Parafilm® and stored at 4°C.

Methanol anhydrous:
Methanol (99.8%) was diluted to a final concentration of 90% using double-distilled water. The solution was stored at RT.

HEK (human embryonic kidney) 293 cell medium:
HEK-293 cell medium was made by combining DMEM/F-12 with GlutaMAX™ Supplement and 10% fetal bovine serum (FBS). The medium was stored at 4°C for up to one month.

Paraformaldehyde:
Paraformaldehyde was diluted with 1x D-PBS to a final concentration of 4%. The solution was stored protected from light for 2 weeks at 4°C.

PBS/EDTA solution:
Ethylenediaminetetraacetic acid (EDTA, 0.5 M) was diluted using 1x D-PBS to a final concentration of 0.5 mM. The solution was stored at RT.

Puromycin dihydrochloride:
Puromycin dihydrochloride was diluted to a concentration of 200 mg/µL using double-distilled water. The solution was aliquoted and stored at -20°C for up to six months.

Rock-Inhibitor Y-27632:
Rock-Inhibitor was diluted to a concentration of 10 mM using 1x D-PBS, aliquoted and stored for six months at -20°C. Before use the solution was further diluted to a final concentration of 10 µM. The final concentration was stored at 4°C up to 14 days.

RPMI 1640 + B27:
RPMI 1640 + B27 medium was created by adding 10 mL of B27™ Plus Supplement to 500 ml of RPMI Medium 1640 L-Glutamine. The solution was stored at 4°C for up to four weeks.

RPMI 1640 no glucose + B27:
RPMI 1640 no glucose + B27 medium was created by adding 10 mL of B27™ Plus Supplement to 500 ml of RPMI 1640 Medium without glucose. The solution was stored at 4°C for up to four weeks.

RPMI 1640 + B27-insulin:
RPMI 1640 + B27-insulin medium was created by adding 10 mL of B27™ Supplement minus insulin to 500 mL of RPMI Medium 1640 L-Glutamine. The solution was stored at 4°C for up to four weeks.

Triton™ X-100 solution:
Triton™ X-100 was diluted to a final concentration of 0.2% using 1x D-PBS. The solution was stored at RT.

Trypan Blue solution:
Trypan blue (0.4%) was diluted to a final concentration of 0.2% using 1x D-PBS. Aliquots were stored at RT.

VF2.1.Cl solution:
VF2.1.Cl solution was prepared by mixing 1 µL of 2 mM VF2.1.Cl with 1µL of Pluronic® F-127. This mixture was added to 5 mL of FluoroBrite™ DMEM containing Hoechst 33258 at a concentration of 4 µg/mL.

Wnt-C59:
Wnt-C59 was diluted with DMSO to a final concentration of 2 mM, aliquoted and stored at -80°C for up to six months.

3.1.8 Cell lines and bacteria

In this study two different cell lines were used:

1) HEK (human embryonic kidney) 293 cells were purchased online from Merck KGaA, (Darmstadt, Germany, #85120602).
2) The human induced pluripotent stem cell line LMNA was a gift from Joseph Wu (Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, USA). The characteristics of the line are already published (Sun, Yazawa et al. 2012).

For plasmid amplification One Shot® TOP10 chemically competent *E.coli* (Invitrogen™, ThermoFisher, # C404003) were used.

3.1.9 Technical devices

Table 10 lists all the technical devices used.

Table 10: Devices

Device	Manufacturer
Andromeda Vacuum Medical Sterilization Autoclave, #N-AUTOCL0001	Alternup Medical, Pontcharra-sur-Turdine, France
BD FACSAria™ III	BD Biosciences, San Jose, CA
C1000 Touch Thermal Cycler,	Bio-Rad Laboratories, Inc., Munich, Germany
Carl Roth™ Microwave Oven	Carl Roth™, Karlsruhe, Germany
Electric Heated Water Bath Laboratory Circulating Water Bath	BioBase, Wolfenbüttel, Germany
Eppendorf Research® plus pipettes	Eppendorf, Hamburg, Germany
ES Series Lab Refrigerators	ThermoFisher Scientific, Waltham, MA
Gel Doc EZ Gel Documentation System	Bio-Rad Laboratories, Inc., Munich, Germany
Heidolph Polymax	Heidolph Instruments GmbH & CO. KG, Schwabach, Germany
Heraeus™ Megafuge™ 8 Small Benchtop Centrifuge	ThermoFisher Scientific, Waltham, MA
Herasafe 2030i Class II Biological Safety Cabinets	ThermoFisher Scientific, Waltham, MA
Kinetic image cytometer IC200	Vala Sciences, San Diego, CA
LAB Scale PCE-BS 300, #PCE-BS 300	PCE Instruments™, Alicante, Spain
Leica SP5 upright Confocal, multi-photon, FLIM	Leica Camera, Wetzlar, Germany
Liquid Nitrogen Tank	Worthington Industries, Columbus, OH
Mr. Frosty™ Freezing Container, #5100-0001	ThermoFisher Scientific, Waltham, MA
NanoDrop 2000c Spectrophotometer	ThermoFisher Scientific, Waltham, MA
New Brunswick™ Innova® 44/44R, Incubator shaker	Eppendorf, Hamburg, Germany
Nucleofector™II	Amaxa Biosystems, Lonza Cologne, Cologne, Germany
N-Storm, Super-Resolution Microscopy	Nikon, Tokio, Japan
Orion Star A321 pH Portable Meter	ThermoFisher Scientific, Waltham, MA
CFX96 Touch Real-Time PCR Detection System	Bio-Rad Laboratories, Inc., Munich, Germany
Sub-Cell GT Electrophoresis Cell	Bio-Rad Laboratories, Inc., Munich, Germany
TC20™ Automated Cell Counter	Bio-Rad Laboratories, Inc., Munich, Germany
Thermo Scientific™ HERAfreeze™ HFU T Series -86°C Upright Ultra-Low Temperature Freezers	ThermoFisher Scientific, Waltham, MA
TSX Series High-Performance Lab Freezer	ThermoFisher Scientific, Waltham, MA
Vortex-Genie 2	Scientific Industries, Bohemia, NY
Wizard Mini-Fuge	Scientific Industries, Bohemia, NY
ZEISS Axio Observer	Carl Zeiss Microscopy, LLC, White Plains, NY

3.1.10 Websites and software

The following websites and software were used during the course of this study.

Basic Local Alignment Search Tool (BLAST), https://blast.ncbi.nlm.nih.gov

Benchling, Cloud-Based Informatics Platform for Life Sciences, https://benchling.com

BioRender, https://biorender.com/

CyteSeer, Vala Sciences, http://www.valasciences.com/instruments/cyteseer

FlowJo, https://www.flowjo.com

Image J, https://imagej.net

National Center for Biotechnology Information (NCBI), https://www.ncbi.nlm.nih.gov

Tide, https://tide.nki.nl

3.2 Methods

3.2.1 Cell culture

Human induced pluripotent stem cells:

Human iPS cells were maintained using E8 medium with a daily medium change. The cells were cultured on Matrigel Matrix coated plates. Corning Matrigel Matrix was thawed overnight at 4°C, aliquoted based on the manufactures instructions in order to achieve the specific protein concentration required and stored at -20°C. The plates were prepared by thawing one aliquot and diluting it with the required amount of DMEM/F-12 medium. The solution was pipetted into the wells yielding a concentration of 9 µg/cm^2 Matrigel (Burridge, Matsa et al. 2014). The plates were swirled and kept in the 37°C incubator for 30 min before further use or stored for up to 4 weeks.

The cells needed to be passaged every 5 to 7 days in order to prevent overgrowth and spontaneous differentiation. Passaging was performed by aspirating the E8 medium, and adding 1 mL of 0.5 mM PBS/EDTA solution for 5 to 10 minutes. The plate was kept in the 37°C incubator. Meanwhile a Matrigel-coated 6-well plate was prepared by aspirating the medium off the wells, and adding 1 mL of E8-Y medium per well. A 15 mL Falcon tube was prepared by adding 11 mL of E8-Y. The iPS cells were taken from the incubator, PBS/EDTA was aspirated, and the cells were detached from the wells by blasting 1 mL of E8-Y medium against the cell surface. The dissociated cells were added to the 11 mL E8-Y, and the Falcon tube was inverted gently to guarantee equal cell dispersion. The cells were plated out by adding 1 mL of the solution to each well of a 6-well plate. On the next morning the medium was changed to E8, and the cells were maintained by daily E8 medium change.

HEK-293 cells:

HEK-293 cells were maintained using HEK-293 cell medium, and cultured on non-coated plates. The medium was changed every other day. Cells were passaged as soon as they reached 70-80% confluency by adding 2 mL TrypLE to each 6-well plate. After incubation at 37°C for 5 min, the cell solution was transferred into a 15 mL Falcon tube, and the TrypLE was inactivated by adding 10 mL of HEK-293 cell medium. Cells were dissociated by gentle pipetting and reseeded back into a 6-well plate by adding 2 mL to each well.

3.2.2 Cardiomyocyte differentiation

Cardiomyocyte differentiation was performed according to the protocol of Lian *et al.* (Lian, Zhang et al. 2013). Figure 5 gives an overview of the differentiation protocol.

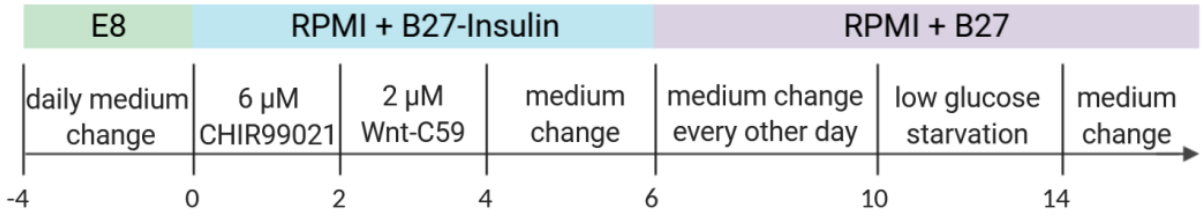


Figure 5: Cardiomyocyte differentiation protocol.

As described above human iPSC culture was maintained using E8 medium. As soon as the human iPSCs achieved 70-80% confluency the cardiomyocyte differentiation protocol was started. The differentiation protocol is based on regulation of the Wnt signaling pathway, and yields up to 80% of cardiomyocytes (Lian, Zhang et al. 2013). Differentiation was started by adding 6 µM of CHIR99021, a Wnt signaling pathway activator. After two days the medium was changed, and 2 µM of Wnt-C59, a Wnt signaling pathway inhibitor, was added. At day 4 (d4) the medium was changed to pure RPMI 1640 + B27-Insulin. After further two days RPMI 1640 + B27 was used for cell maintenance. In order to purify the cell population glucose starvation using RPMI 1640 no glucose + B27 was performed between d10 and d14 depending on the occurrence of beating cardiomyocytes (Tohyama, Hattori et al. 2013). Afterwards, the medium was replaced every second day using RPMI 1640 + B27. Typically, first beating cells could be detected by d10 to d12.

3.2.3 Plasmid amplification

For plasmid amplification One Shot® TOP10 chemically competent *E.coli* were used. After thawing the bacteria on ice, about 20-30 ng of plasmid were added to 1 vial of One Shot® TOP10 chemically competent *E.coli* (50 µL) and kept for 5 min on ice. Afterwards, the vial was heated to 42°C for 2 min and thereafter 250 µL of SOC media were added. The mixture was kept in a heated shaker (37°C) for one hour with continuous shaking (1,000 rpm). Next, 50 to 100 µL of this solution were added onto a LB agar petri dish with ampicillin (see 3.1.7), and put for 16 to 18 hours into a 37°C incubator. On the next day, several clones were picked, and each clone was individually transferred into 5 mL LB inoculated with 5 µL ampicillin (100 mg/mL). The Falcon tubes were placed into a 37°C incubator for about 16-18 hours. Next, the solution was scaled up by adding 100 µL of the first solution into 25 mL LB inoculated with 25 µL ampicillin (100 mg/mL) and left again in the 37°C incubator for 16-18 hours. In order to isolate the plasmids, the QIAGEN Plasmid Midi Kit was used as recommended by the manufacturer. In brief, plasmid purification relies on alkaline lysate. The lysate was centrifuged through a membrane that selectively binds plasmids, and removes impurities. After a washing step the plasmid was eluted in nuclease free water. Plasmids were stored at -20°C until further use. In addition, glycerol stocks were produced by adding 500 µL of the liquid bacterial culture to 500 µL of a glycerin and nuclease free water solution (50:50). The glycerol stocks were stored at -80°C.

3.2.4 DNA and RNA isolation

Isolation of genomic DNA and total RNA from HEK293 and iPSCs were performed using DNeasy Blood & Tissue Kit and RNeasy Mini Kit according to the manufacturer's instruction. In brief, DNA isolation was performed by lysing the cells using proteinase K. The lysate was centrifuged through a membrane that allows contaminants to pass, but selectively binds DNA. After two washing steps the DNA is then eluted in nuclease free water. The isolated DNA was stored at -20°C.

RNA isolation was performed by lysis and homogenization of the cells with a simultaneous inactivation of RNases in order to prevent RNA degradation. The lysate was pressed through a membrane allowing the removal of genomic DNA with the addition of DNase for further

DNA digestion. Ethanol was added and the lysate was centrifuged through a membrane binding the RNA and letting contaminants pass. After a washing step the RNA is eluted in nuclease free water. The isolated RNA was stored at -80°C.

To assess quantity and purity of the DNA and RNA, spectrophotometric analysis using the NanoDrop™ was performed. An A260/280 ratio of approximately 1.8 ensured DNA purity. An A260/280 ratio of approximately 2.0 ensured RNA purity.

3.2.5 cDNA production

Conversion of RNA to cDNA was performed using the High Capacity cDNA Reverse Transcription Kit according to the manufacturer's instructions. First, a master mix was prepared on ice as described in table 11.

Table 11: Components for master mix

Components	**Volume [µL]**
10x RT buffer	2.0
25X dNTP Mix (100 mM)	0.8
10X RT random primers	2.0
MultiScribe™ ReverseTranscriptase	1.0
RNase Inhibitor	1.0
Nuclease-free H_2O	3.2
Total per reaction	10.0

After mixing the master mix gently 2 µg of RNA were added and the reverse transcription was performed using the C1000 Touch Thermal Cycler according to conditions specified in table 12. The synthesized cDNA was stored at -20°C.

Table 12: Settings for reverse transcription

	Temperature [°C]	**Time [min]**
Step 1	25	10
Step 2	37	120
Step 3	85	5
Step 4	4	indefinite

3.2.6 Polymerase chain reaction

Polymerase chain reaction (PCR) allows the amplification of DNA segments by thermal cycling. The primers used to amplify the regions of interest are listed under 3.1.5. The thermostable DNA polymerase used for the assembly of new DNA strands is provided with the GoTaq® DNA Polymerase kit, and used according to the manufacturer's instructions. Table 13 lists all components necessary for performing a PCR using the C1000 Touch Thermal Cycler as specified in table 14.

Table 13: Components for PCR

Components	Final concentration
GoTaq® DNA Polymerase	1.25 U
GoTaq® Reaction buffer	1x 1.5 mM $MgCl_2$
PCR nucleotide mix	10 mM
Upstream primer	1 µM
Downstream primer	1 µM
Template DNA	100 ng
Nuclease free water	fill up to 50 µL

Table 14: Settings for the C1000 Touch Thermal Cycler

Step	Time	Temperature [°C]	Number of cycles
Initial Denaturation and Polymerase activation	2 minutes	95	1 cycle
Amplification			25-35 cycles
Denaturation	30 seconds	94	
Annealing	40 seconds	58	
Extension	40 seconds	72	
Final Extension	5 minutes	72	1 cycle
Cooling	indefinite	4	1 cycle

3.2.7 Quantitative real-time polymerase chain reaction (qRT-PCR)

Quantitative real-time polymerase chain reaction (qRT-PCR) allows the assessment of gene expression based on the principles of PCR, using cDNA as a template. This technique allows the relative quantification of the amplified gene products with the help of fluorochromes. For relative quantification *Beta-Actin (ACTB)* was used as internal control. In this study non-specific detection of double-stranded DNA was achieved using the Power SYBR™ Green PCR Master Mix kit. After every amplification cycle the amount of fluorescence was recorded. The cycle when the fluorescence increases above the background threshold is a measure to

assess the quantity of the amplified gene fragments. In order to assess the specificity of the detected fluorescence, after the amplification steps a melting curve analysis is performed. By slow increase of the temperature the created double-strand DNA fragments separate at a specific point allowing to distinguish between incorrect aggregation, primer dimers and the amplified double-strand DNA fragments. Table 15 lists the components used for performing qRT-PCR. Table 16 shows the settings using the CFX96 Touch Real-Time PCR Detection System.

Table 15: Components for qRT-PCR

Components	**Quantity**
Power SYBR Green PCR Master Mix (2x)	10 µL
Forward primer	100 nM
Reverse primer	100 nM
Template DNA	100 ng
Nuclease free water	fill up to 20 µL

Table 16: Settings for the CFX96 Touch Real-Time PCR Detection System

Step	**Time**	**Temperature [°C}**	**Number of cycles**
Uracil DNA glycosylase activation	2 minutes	50	Hold
Polymerase activation	2 minutes	95	Hold
Amplification			40 cycles
Denaturation	15 seconds	95	
Anneal/Extend	60 seconds	60	
Melt curve stage			1 cycle
Step 1	15 seconds	95	
Step 2	60 seconds	60	
Step 3	15 seconds	95	
Cooling	Indefinite	4	1 cycle

Data analysis was performed in accordance to the Pfaffl method (Pfaffl 2004):

$$\textit{Gene expression ratio} = \frac{(E\ gene\ of\ interest)^{\Delta Ct\ gene\ of\ interest}}{(E\ housekeeping\ gene)^{\Delta Ct\ housekeeping\ gene}}$$

E is the equation that refers to the primer efficiency. In order to assess primer efficiency a standard curve is generated by serial dilution (1:10) with at least 5 points. The measured Ct values are used to create the slope of the line:

Ct= slope x log(sample quantity) + intercept y-axis

The slope of the line allows the calculation of the primer efficiency:

$$E\ (\%) = 10^{\left(\frac{-1}{slope}\right)} - 1 \times 100$$

For the Pfaffl method the converted primer efficiency was used and calculated as follows:

$$Converted\ primer\ efficiency = \left(\frac{(Primer\ efficiency\ (\%)}{100}\right) + 1$$

3.2.8 Sequencing

Sequencing was performed using the services from the Protein and Nucleic Acid Facility at Stanford School of Medicine. Before sequencing the PCR products were purified using the QIAquick PCR Purification Kit according to the manufacturer's instruction. In brief, the PCR products were filtered through a silica membrane binding DNA, while other contaminants pass through. After two more washing steps, the absorbed DNA was eluted in nuclease free water. Afterwards, the samples were prepared as described in table 17.

Table 17: Sample preparation for sequencing

Component	**Volume [µL]**	**Concentration**
Purified DNA template	10	Plasmids: 100 ng/µL PCR product: 10 ng/µL
Primer	2	5 µM

3.2.9 Gel-electrophoresis

For gel electrophoresis 1% agarose gel was prepared as follows: 0.5 g of agarose was dissolved by microwaving in 50 ml of Tris acetate-EDTA buffer. Afterwards, 10 µL of a 50 µg/mL ethidium bromide solution was added to every 50 mL of the agarose solution. The solution was poured into the appropriate gel chamber and left at RT to harden.

One microliter of each amplified DNA sample was diluted with 19 µL of nuclease free water and 4 µL of 6x DNA gel loading dye was added, and the solutions were pipetted into the

pockets. As a marker the 1 kb plus DNA ladder from Invitrogen by Thermo Fisher Scientific was used. The gels were loaded into the electrophoresis cells filled with 1x Tris acetate-EDTA buffer for 40 min at 120 V. Imaging was performed using the Gel Doc EZ Gel Documentation System.

3.2.10 Immunocytochemistry

Immunocytochemistry (ICC), based on the indirect method, was used to identify certain proteins of interest in human iPSCs and iPSC-derived cardiomyocytes. For the indirect method two different antibodies are needed. First, the unlabeled primary antibody that is specific to the antigen of interest. Second, the fluorophore-labeled secondary antibody that binds to the invariant part of the primary antibody. Table 4 and 5 list all the antibodies used in this study.

In order to prepare the cells for ICC, they were transferred from a 6-well into 8-well chamber slides. Therefore, the culture media was removed, and the cells were detached using 2 mL of Accutase. After incubation for 2 minutes at 37°C, the Accutase was deactivated by adding 2 mL of the appropriate culture media: E8 medium for iPSCs and RPMI 1640 + B27 for iPSC-derived cardiomyocytes. The cells were counted by taking 10 µL of the cell solution and adding 10µL of Trypan Blue Stain in a separate Eppendorf tube. The solution was pipetted into a Countess slide and cell count was established using the TC20™ Automated Cell Counter. Meanwhile, the cell solution was centrifuged for 5 minutes at 300 x *g*. The supernatant was discarded, and the pellet was resuspended using the appropriate culture media (E8 for iPSCs and RPMI 1640 + B27 for iPSC-derived cardiomyocytes) to obtain a concentration of 50 thousand cells per mL. The cells were replated into the 8-well chamber slide with about ten thousand cells per well (200 µL). After 24 hours the culture media was removed and the cells were fixed with 4% paraformaldehyde for 5 minutes at RT. After washing the cells with 1x D-PBS (3-times), they were permeabilized with 0.2% Triton™ X-100 solution for 5 minutes. Unspecific binding sites were blocked with a treatment of 1x D-PBS containing 3% BSA for 30 minutes. In the next step the primary antibody was added and the chamber slides were stored at 4°C for 16-18 hours. After washing the cells with 1x D-PBS (3-times), the secondary antibody was added and stored protected from light for 2 hours at RT.

After a final washing process (3-times), the walls of the chamber slides were removed, kept protected from light for 20 minutes to dry. Afterwards, Mounting Medium with DAPI was added, and the slides were covered with cover glass.

Imaging was performed using either N-Storm Super-Resolution Microscopy or Leica SP5 upright Confocal.

3.2.11 Flow cytometry

Flow cytometry allows analyzing or sorting single cells based on their characteristics including cell size, structure and shape, as well as fluorescence. Among others, for this study the proteins of interest were labeled with antibodies listed in table 4 and 5, allowing a fluorescence-activated cell analysis. The cells were prepared as follows:

1) Cells were washed with 1x D-PBS and treated with 1 mL of Accutase for iPSC-derived cardiomyocytes for 2 min or 1 mL of Trypsin-EDTA for HEK-293 for 5 min. During incubation the cells were stored at 37°C for dissociation.
2) Dissociated cells were transferred into a 15 mL Falcon tube, topped up with 9 mL of 1x D-PBS and centrifuged for 4 min at 300 x *g*.
3) Further preparation for HEK-293 cells was as follows:
    a. The supernatant was removed, and the cells were resuspended in 500 µL of D-PBS containing 2% BSA.
    b. The cells were filtered into round-bottom tubes with cells strainer caps and placed on ice until sorting.
4) Further preparations for hiPSC-derived cardiomyocytes were as follows:
    a. The supernatant was removed, and the cells were resuspended in 1 mL of 4% paraformaldehyde. The solution was left at RT for 20 minutes, and afterwards centrifuged for 4 minutes at 300 x *g*.
    b. The supernatant was removed. The cells were resuspended in 90% ice-cold methanol, left for 15 minutes at 4°C, and centrifuged again for 4 minutes at 300 x *g*. The supernatant was removed.

c. The pellet was resuspended in 2 mL of D-PBS containing 4% BSA, left for 5 minutes at RT, and centrifuged again for 4 minutes at 300 x *g*. This step was repeated once.
d. The pellet was resuspended in D-PBS containing 4% BSA and 0.2% Triton™ X-100, the primary antibody (diluted in D-PBS containing 3% BSA; see 3.1.4) was added for 60 minutes, and kept at RT. The suspension was centrifuged for 4 minutes at 300 x *g*, resuspended in D-PBS containing 4% BSA and 0.2% Triton™ X-100, and centrifuged again for 4 minutes at 300 x *g*.
e. The pellet was resuspended in D-PBS containing 4% BSA with 0.2% Triton™ X-100, the secondary antibody (see 3.1.4) was added for 30 minutes, and kept at RT in the dark. The solution was centrifuged for 4 minutes at 300 g, resuspended in D-PBS containing 4% BSA and 0.2% Triton™ X-100, and centrifuged again for 4 minutes at 300 x *g*.
f. The pellet was resuspended in D-PBS containing 4 % BSA und kept on ice for flow cytometry.

The flow cytometry was performed using BD FACSAria™ III.

3.2.12 Karyotyping

Karyotyping was performed by using the service of WiCell Research Institute (Madison, WI). The human iPSCs were seeded in a T25 flask, provided by the company, and shipped as live cells heated by warmed cool packs. After 7 to 10 days the company provided a full characterization and interpretation of the karyogram.

3.2.13 CRISPR/Cas9

Genome editing was performed utilizing the CRISPR/Cas9 technology. Figure 6 gives an overview of the intended strategy and schematic presentation of the construct.

The design of the donor construct pursues to delete the stop codon of the *MYL2* gene in order to allow the simultaneous transcription of the tdTomato fluorescence gene. MYL2 and

tdTomato are linked using a P2A sequence. P2A allows the cleavage of the two genes, generating two separate proteins. Furthermore, the construct contains a Puromycin resistance cassette, allowing an antibiotic selection of cells with successful integration of the donor construct. This section is flanked with loxP sites ensuring the possibility of excision of the antibiotic resistance cassette. The donor construct is equipped with a left and right homology arm according to the genomic DNA sequence 500 base pairs left and right of the cutting site.

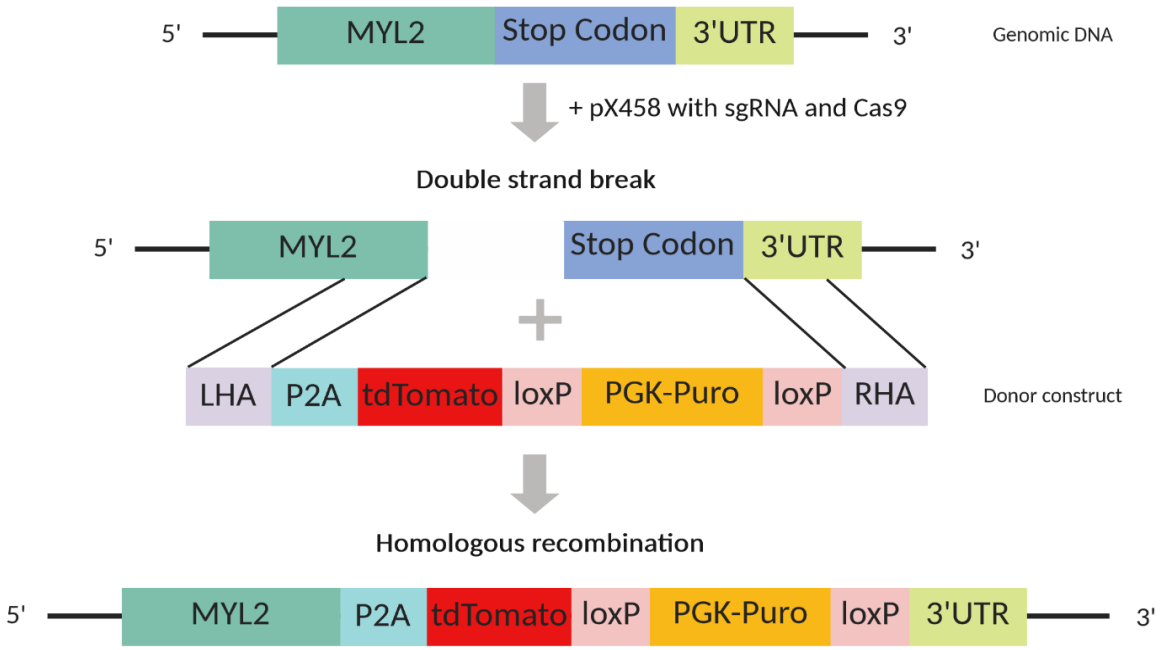


Figure 6: Flow chart CRISPR/Cas9 genome editing of the *MYL2* locus. UTR, untranslated region; LHA, left homology arm; loxP, locus of X-over P1; PKG, phosphoglycerate kinase 1 promoter; Puro, Puromycin

In order to achieve precise genome editing using CRISPR/Cas9 the selection of the appropriate sgRNA is crucial. For sgRNA design the Benchling software was used (www.benchling.com). Three different sgRNAs located close to the insertion site were selected (Table 8) and cloned into the pX458 vector (Fig. 4).

First, the oligonucleotides were prepared as described in table 18 and annealed using the settings described in table 19.

Table 18: Components for annealing the sgRNA oligonucleotides

Component	Quantity [µL]
sgRNA upper strand (100 µM)	1
sgRNA lower strand (100 µM)	1
10x T4 ligation buffer	1
T4 Polynucleotide Kinase	1
Nuclease free water	6

Table 19: Thermocycler setting for oligonucleotide annealing

Temperature	Time [min]
37°C	30
95°C	5
95°C ramped down to 25°C	Ramp: 5°C/min

The final product of the annealed oligonucleotides was diluted 200-fold using nuclease free water, and the digestion-ligation reaction was performed as shown in table 20.

Table 20: Digestion-ligation reaction

Component	Quantity
pX458 plasmid	100 ng
Annealed oligonucleotides	2 µL
10x Tango buffer	2 µL
DTT	1 mM
ATP	1 mM
Fast Digest Bpil (Bbsl; 10 U/µL)	1 µL
T7 DNA ligase (3,000,000 U/mL)	0.5 µL
Nuclease free water	Fill up to 20 µL

The mixture was incubated using the Thermocycler as described in table 21.

Table 21: Thermocycler settings for digestion-ligation reaction

Step	Time	Temperature	Number of cycles
Digestion-ligation			6 cycles
	5 min	37°C	
	5 min	23°C	
Cooling	Indefinite	4°C	1 cycle

As final step, the product was treated with Plasmid-Safe ATP-Dependent DNase in order to clean up for linear double-stranded DNA and linear closed-circular single-stranded DNA. Table 22 shows the approach for the treatment with Plasmid-Safe ATP-Dependent DNase. The solution was incubated for 30 minutes at 37°C.

Table 22: Clean up with Plasmid-Safe ATP-Dependent DNase

Component	Quantity [µL]
Digestion-ligation product	11
10x PlasmidSafe buffer	1.5
10 mM ATP	1.5
PlasmidSafe exonuclease	1

The final plasmid was amplified via bacterial transformation as described above (see 3.2.3). For each sgRNA two different clones were picked and sequenced to verify the correct insertion of the sgRNA oligonucleotides using the U6 primer. Sequencing was performed as described above (see 3.2.8).

In order to quantify the cutting efficiency of the selected sgRNAs, HEK-293 cells were transfected with the pX458 plasmid containing the sgRNA1, sgRNA2 and sgRNA3, respectively. Also, a mock transfection without any plasmid was included. Transfection was performed using the Nucleofector™II device. When reaching 70-80% confluency, HEK-293 cells were nucleofected as described below (see 3.2.14).

After nucleofection the HEK-293 cells were cultured for 24 hours without any media change. Next, the cells were prepared for flow cytometry as described above (see 3.2.11). GFP positive cells were sorted, cultured for 2-3 more days, and DNA extraction was performed as described above (see 3.2.4). In order to prepare the samples for testing the cutting efficiency

of the sgRNAs, a PCR (see 3.2.6) spanning the expected cutting region was performed by using the primer pair "sgRNA cutting efficiency" (Table 6). Purified PCR products from HEK293 cells transfected with sgRNAs and untransfected control cells were sent in for sequencing (see 3.2.8). The sequencing results were analyzed using the TIDE software (https://tide.nki.nl), which allows the quantification of the cutting efficiency.

After the establishment of a functional sgRNA in the pX458 plasmid, and the design of the donor template (see 3.1.6), nucleofection of human iPSCs was performed for genome editing (see 3.2.14). Twenty-four hours later, the cells were treated with 0.2 µg/mL puromycin selecting individual clones harboring the donor construct. Clone picking was performed under the microscope by scraping off the individual clones and replating each clone into a Matrigel-coated 24-well plate. Special attention was devoted to not mixing the individual clones. As soon as the clones reached 70-80% confluence, they were replated into duplicate Matrigel-coated 6-well plates. One well was used for DNA isolation (see 3.2.4), cells in the second well were kept for further culture. After DNA isolation a PCR using the primer pair "Donor sequence insertion in MYL-2 locus" (Table 6) was performed to check for the correct insertion of the donor template (see 3.2.6). Clones in which PCR products of the correct size in the gel-electrophoresis were seen (see 3.2.9), were sent out for sequencing (see 3.2.8) using the forward primer of the PCR reaction.

3.2.14 Nucleofection

Nucleofection was performed using HEK-293 or human iPS cells as described below. The cells were cultured in different media as described above (see 3.2.a Cell culture). Otherwise the protocol did not vary between the different cell types.

In preparation for nucleofection Accutase and the different culture media were prepared by adding Blebbistatin at a 1:4000 dilution. Additionally, a different batch of culture media was prepared by adding Thiazovidin at a 1:5000 dilution. For nucleofection the commercially available Human Stem Cell Nucleofector® Kit1 was used. For each reaction 82 µL of the provided P3 solution, 18 µL of the provided supplement 1, and 5 µg of the pX458 plasmid were mixed into an Eppendorf tube. Additionally, for the nucleofection of iPSCs 1 µg of pUC57-MYL2 donor vector were added. The cells were detached using Accutase with

Blebbistatin. After incubation for 2 minutes at 37°C, the Accutase was deactivated by adding the appropriate culture media with Blebbistatin. The cells were gently dissociated into single cells by pipetting up and down, and collected into a 15 mL Falcon tube. The cells were counted by taking 10 µL of the cell solution and adding 10µL of Trypan Blue Stain in a separate Eppendorf tube. The solution was pipetted into a Countess slide and cell count was established using the TC20™ Automated Cell Counter. Meanwhile, the cell solution was centrifuged for 5 minutes at 300 x *g*. The supernatant was discarded, and the pellet was resuspended using the appropriate culture media with Blebbistatin to obtain a concentration of one million cells per mL. Afterwards, 1 mL of the cell solution was added into a 15 mL Falcon tube and centrifuged again for 3 min at 300 x *g* The supernatant was removed, and the cells were resuspended in the above mentioned solution containing P3, supplement 1, the pX458 plasmid, and for iPSCs the pUC57-MYL2 donor vector. The mixture was transferred into the provided cuvettes, and nucleofected using the Nucleofector™II. Immediately afterwards, 1 mL of the appropriate culture medium with Thiazovidin was added into the cuvettes, and the cells were allowed to rest for 3-4 minutes at RT. Next, the cells were plated into a Matrigel-coated 6-well plate, and 2 mL of additional culture medium with Thiazovidin were added. The cells were transferred into the 37°C incubator, and medium was changed after 24 to 48 hours.

3.2.15 Electrophysiological analysis

Electrophysiological analysis was performed as described by McKeithan *et al*. (McKeithan, Savchenko et al. 2017) in collaboration with the Mercola laboratory, Stanford Medicine. Action potential (AP) kinetics were measured fully automated by using a small molecule fluorescent voltage sensing probe (Fluovolt™; VF2.1.Cl) (Figure 7).

The cells were prepared as follows: First, human iPSC-derived CMs were passaged on d15 of cardiac differentiation into Matrigel-coated 384-well plates with about 50 thousand cells per well. The cells were cultured for another 15 days with daily media change before AP kinetics were measured. During cell preparation, they were placed on a 37°C dry heat block to avoid temperature fluctuation. The used media were warmed to 37°C in a water bath. The cells were washed five times by removing 50 µL of the cell culture media, and replacing it with 50

µL of FluoroBrite™ DMEM. Afterwards, 50 µL of VF2.1.Cl solution were added, and the plate was placed into the incubator for 50 minutes at 37°C. Next, the cells were washed four times using FluoroBrite™ DMEM, and the plate was kept again in the incubator for 10 minutes. Prior to imaging, 50 µL of media were removed and, 50 µL of 100 mM 4-Aminopyridine were added. Imaging was performed using the Kinetic Image Cytometer IC200 at an acquisition frequency of 100 Hz with the excitation wavelength of 485/20 nm, and an emission filter of 525/30 nm using a 0.75 NA 20x Nikon Apo VC objective. Image analysis was performed using the CyteSeer software.

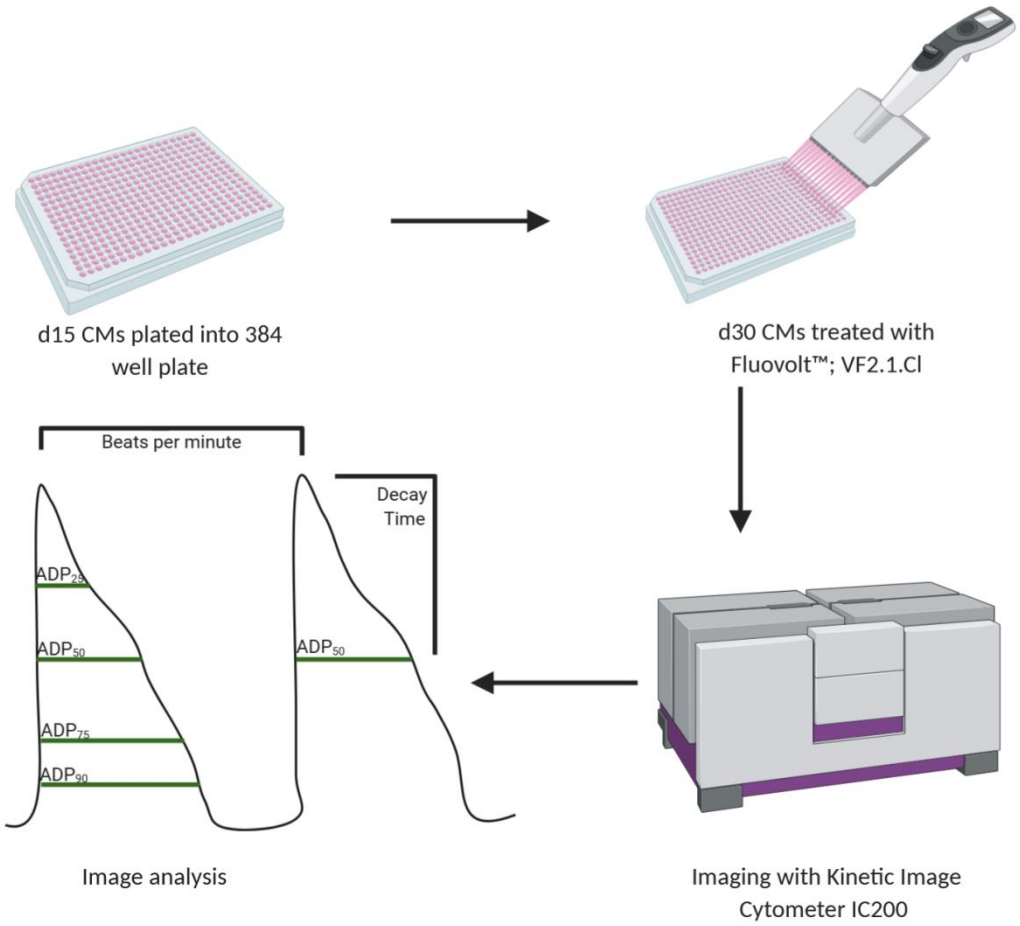


Figure 7: Schematic illustration of the electrophysiological analysis. After transferring hiPSC-derived CMs into a 384 well plate, the cells were treated with Fluovolt™ after 15 days, imaged and data analysis was performed. ADP, action potential duration.

3.2.16 Traction force microscopy

Traction force analysis was performed as described by Ribeiro *et al.* (Ribeiro, Schwab et al. 2017) in collaboration with the Pruitt laboratory, Stanford University. For this purpose, the

human iPSC-derived CMs were cultured as single cells on polyacrylamide hydrogels, and videos were taken. Each video was analyzed for substrate displacement, which allowed the calculation of contractile force.

3.2.17 Statistical analysis

Data are presented as mean values ± standard error of the mean (S.E.M.).Statistical analysis comparing the tdTomato positive cells and controls was performed using unpaired parametric Student's t-test. P values <0.05 were considered significant.

4. Results

4.1 Efficiency of the CRISPR/Cas9 mediated genome editing

The three selected sgRNAs were evaluated for their cutting efficiency using the TIDE software. SgRNA2 showed a cutting efficiency of 18 %, and was used for the further genome editing in hiPSCs (Fig. 8).

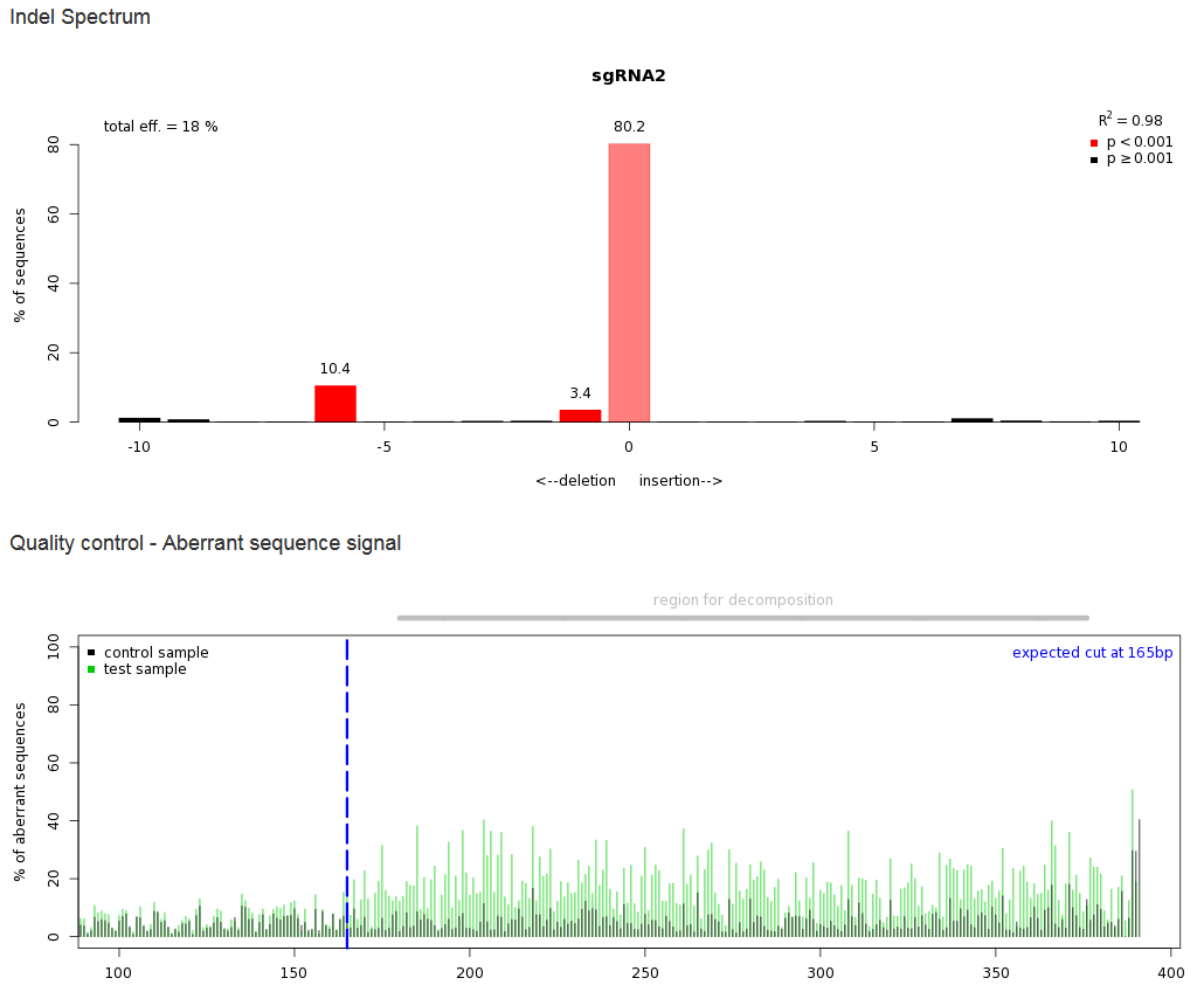


Figure 8: Cutting efficiency of sgRNA2 calculated using the TIDE software. The upper graph depicts the estimated cutting efficiency with 18%. The lower graph shows that the aberrant sequence signal starts at the expected cutting site of the sgRNA.

Genome editing of $1x10^6$ iPSCs resulted in about 150 remaining clones after antibiotic treatment with puromycin. Eighteen individual clones were picked, and each clone was transferred in one well of a 24-well plate, carefully paying attention not to mix cells of different individual clones (Fig 9A). After seven more days, each clone was transferred into two wells of a 6-well plate. One well was kept for further culture; the other was used for DNA extraction, in order to verify the correct integration of the donor sequence. PCR was

performed using primers outside the left homology arm and within the tdTomato sequence (Fig. 9B; primerpair “Donor sequence insertion in *MYL-2* locus”; see 3.1.5). Figure 9C shows that from 14 clones (78%) a PCR fragment of the expected site was amplified, indicating the correct integration of the donor sequence. Four clones (7, 8, 10 and 18) did not show a PCR product, despite an antibiotic resistance to Puromycin. In order to distinguish between heterogeneous or homogenous integration of the donor sequence at the *MYL-2* locus, a separate PCR was performed targeting the wildtype *MYL-2* gene (Primerpair “WT *MYL-2* expression; see 3.1.5). All tested clones displayed a band in the agarose gel electrophoresis, indicating the heterogeneous integration of the donor sequence (Figure 9D). For further verification the PCR products revealing the correct integration of the donor sequence from clones 1 to 5, 9, and 12 to 17 were sequenced. All were showing precise integration (Fig. 9E). Clone 5 was selected for all further experiments and is referred to as MYL-2 reporter line.

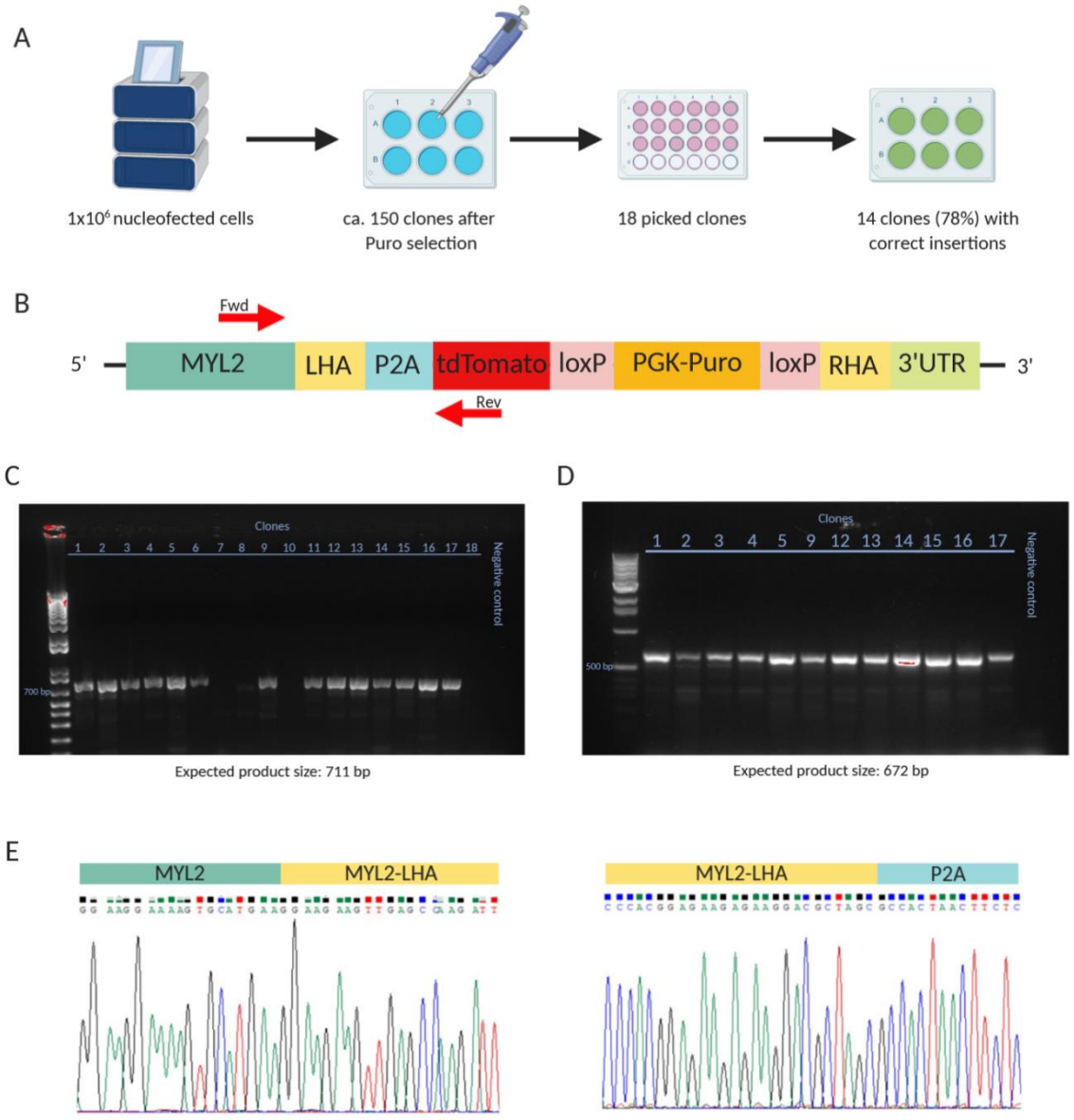


Figure 9: Efficiency of the CRISPR/Cas9 mediated genome editing. (A) Schematic illustration of number of cells/clones after each step after genome editing. **(B)** Scheme of the *MYL2* genomic locus after successful genome editing. Red arrows indicate the location of the used forward (Fwd) and reverse (Rev) primers. **(C)** Detection of correct insertion of the donor sequence in the *MYL2* genomic locus by agarose gel electrophoresis. Clones 1 to 6, 9 and 11 to 17 show bands of the correct size. As a marker the 1kb plus DNA ladder from Invitrogen by Thermo Fisher Scientific was used. **(D)** Detection of the wildtype *MYL2* genomic locus by agarose gel electrophoresis. All tested clones show the wildtype *MLY2* genomic locus. As a marker the 1kb plus DNA ladder from Invitrogen by Thermo Fisher Scientific was used. **(E)** Sanger sequencing of the edited clone 5 showing the correct insertion of the donor sequence.

4.2 Effects of the CRISPR/Cas9 mediated genome editing

One of the biggest concerns regarding genome editing is the occurrence of off-target cleavage sites created by imprecise cutting of the nucleases. Unfortunately, there is no reliable way to assess or truly predict the rate of off-target sites. Lin *et al.* showed that one to three mismatches between genomic DNA and sgRNA can lead to off-target cleavage sites causing indel mutations (Lin, Cradick et al. 2014). Based on this work, all the genomic sequences with up to three mismatches were identified using the software Benchling (Fig. 10A). Three clones (3, 4 and 5) were checked for off-target effects by sequencing. There were no indel mutations to be found at any of the 16 identified locations (Fig. S5).

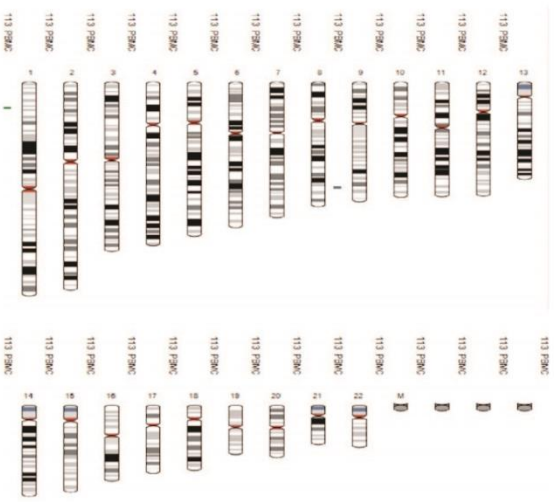
Besides, the generated MYL-2 reporter line (clone 5) was checked for chromosomal aberrations via karyotyping. The cell line showed no chromosomal aberrations after genome editing (Fig. 10B).

Furthermore, the pluripotency status of the edited clones was proven in order to exclude any negative impact by the genome editing technology CRISPR/Cas9. The homeobox protein NANOG is a transcription factor crucial for maintaining pluripotency in stem cells. Its name is derived from Tír na nÓg (Irish for "Land of the Young"), a mythological Celtic land of the young, indicating the ability of pluripotent cells for limitless duplication (Gawlik-Rzemieniewska and Bednarek 2016). TRA-1-81 is an antigen expressed on the transmembrane glycoprotein podocalyxin on the surface of stem cells. Podocalyxin is a sialoglycoprotein and a member of the CD34 family (Schopperle and DeWolf 2007, Heath, Heilbrun et al. 2018). Both, NANOG and TRA-1-81, are commonly used to assess the pluripotent status of stem cells. Figures 10C and 10D show representative stainings of the expression of both pluripotency markers, before and after genome editing for the selected clone.

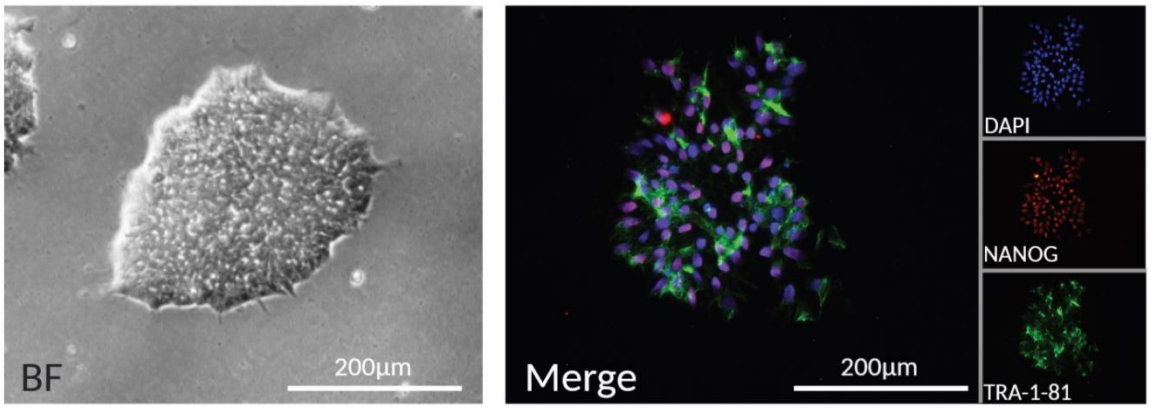
A

Sequence	PAM	Chromosome	Mismatches
CGGAGAAGAGAAGGACTAGG	AGG		
AGGAGAAGAGAAGGACTAGG	TAG	chrX	1
AGGAAAAGATAAGGACTAGG	AAG	chr13	3
TCCAGAAGAGAAGGACTAGG	GAG	chr8	3
AGGAGAAGAAAAGGACTAGC	TGG	chr2	3
CAGAGAAGGCAAGGACTAGG	AGG	chr16	3
TGGACAAGAGAGGGACTAGG	AAG	chr7	3
AGAAGGAGAGAAGGACTAGG	AGG	chr3	3
AGCAGAAGAGAAGGACCAGG	TGG	chr8	3
AGAAGAAGAGAAGGACTAAG	CAG	chr2	3
CTCAGAAGAGAATGACTAGG	GAG	chr3	3
CTGAGAGGAGAGGGACTAGG	AGG	chr14	3
CGGGGATGAGAAGGACAAGG	GAG	chr8	3
CAGAGATGAGAATGACTAGG	CAG	chr10	3
AGGAGAGGAGAAGGGCTAGG	GGG	chr6	3
AGGAGAAGGGAAGGTCTAGG	AAG	chr14	3
CAGAGAAGAGAAAGACCAGG	AGG	chr1	3

B



C



D

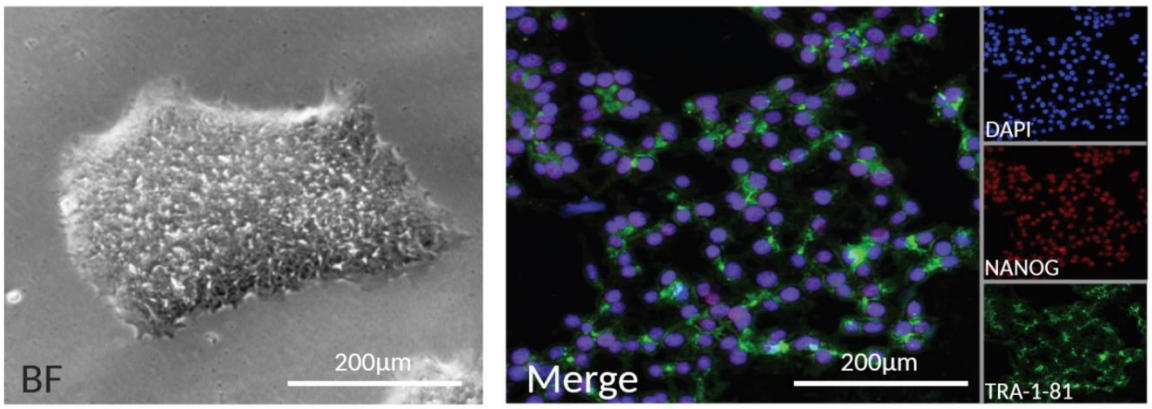


Figure 10: Effects of the CRISPR/Cas9 mediated genome editing. (A) List of potential off-target sites with up to three sgRNA mismatches. **(B)** Karyotype of the selected edited clone 5 showing no chromosomal aberrations. **(C)** Representative images of immunostaining for NANOG and TRA-1-81 of human iPSC line without genome editing. **(D)** Representative images of immunostaining for NANOG and TRA-1-81 of after genome editing. BF: bright field.

4.3 Differentiation of the genome edited iPSC line into cardiomyocytes

The edited cell line, established from the clone 5 was differentiated into cardiomyocytes according to the described protocol by Lian *et al.* (Lian, Zhang et al. 2013) (Fig 11A). At day 8 of differentiation the first beating areas were detectable. Around day 15 a few red cells were visible under the fluorescent microscope. By day 30 of differentiation the majority of cells were red (Fig 11B). To quantify the amount of tdTomato positive cells fluorescence-activated cell sorting (FACS) was performed. Figure 11C shows the rapidly increasing amount of tdTomato positive cells during differentiation, resulting in 55.5% of all cardiac Troponin-T (cTNT) positive cells to express the red fluorescent marker at day 30 of differentiation.

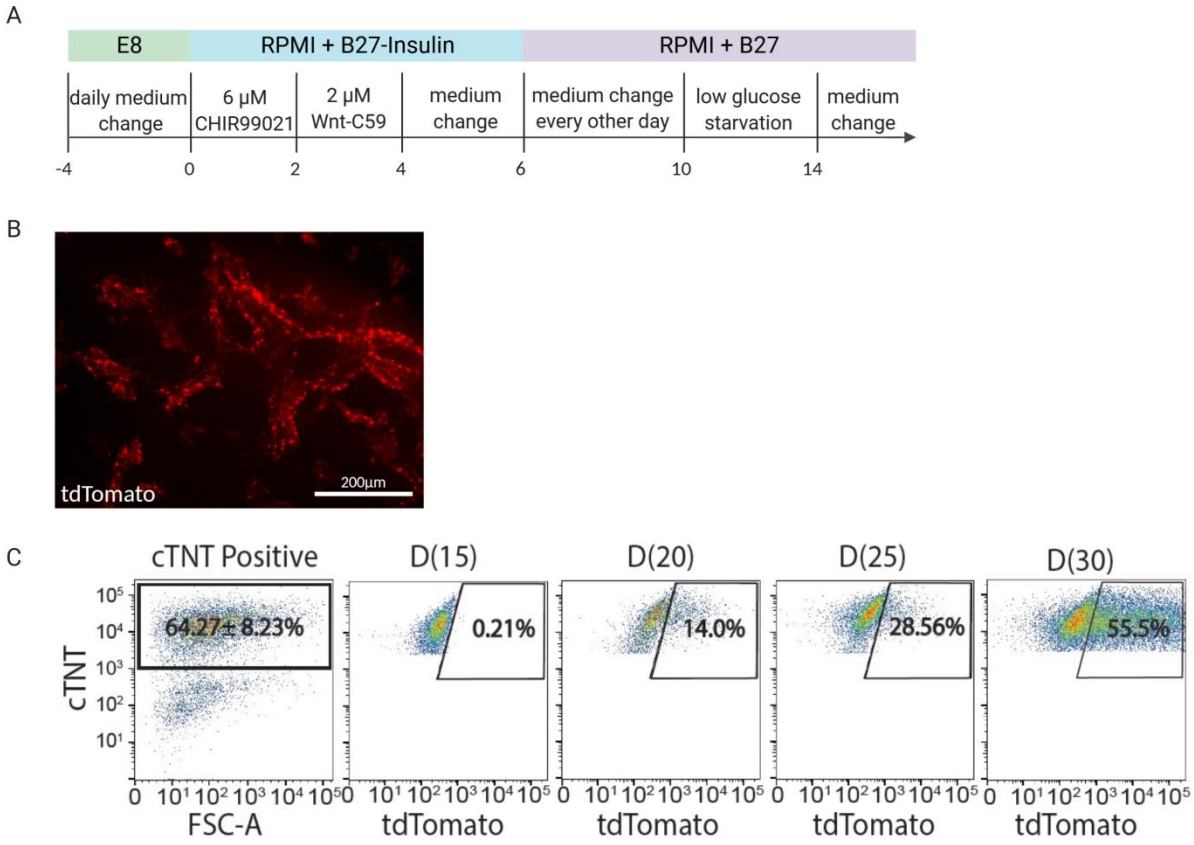


Figure 11: Differentiation of the genome edited iPSC line into cardiomyocytes. (A) Scheme of the cardiomyocyte differentiation protocol according to Lian *et al.* (Lian, Zhang et al. 2013) **(B)** Image of cardiomyocytes (d30) generated from the edited cell line. **(C)** Representative FACS plots from day 15, 20, 25, and 30 cardiomyocytes showing the increase of the tdTomato positive population within the cTNT positive population. The first plot shows the amount of cTNT positive cells within the prepared cells.

4.4 Characterization of the tdTomato positive reporter line

In order to verify that the MYL-2 reporter cell line truly labels ventricular cardiomyocytes, the gene expression pattern is of utter importance. Figure 12A and B show the quantitative real-time PCR (qRT-PCR) results. As described above NANOG is a transcription factor crucial for maintaining pluripotency in stem cells. Octamer-binding transcription factor 4 (OCT-4) is encoded by the *POU5F1* gene and also serves as a transcription factor. It has been shown that it plays a role in maintaining the self-renewal ability of stem cells (Tai, Chang et al. 2005). Both transcription factors are expressed in the pluripotency state of human iPSCs and decrease during cardiac differentiation (Figure 12A). On the other hand the cardiac marker *cTNT* increases during differentiation. More importantly, the results show also an increase of the ventricular cardiomyocytes marker *MYL-2* with a concurrent increase of the linked *tdTomato* (Figure 12B).

In addition, immunostaining of day 32 cardiomyocytes was performed, showing the overlapping expression of MYL-2 and tdTomato (Figure 12C), thereby confirming the linked expression of these two genes after successful genome editing with the CRISPR/Cas9 technology.

Furthermore, cell size quantification was performed on day 30 cardiomyocytes stained for beta-ACTIN and DAPI. Figure 12D shows that tdTomato positive cells (ventricular cardiomyocytes) are significantly larger than tdTomato negative cells.

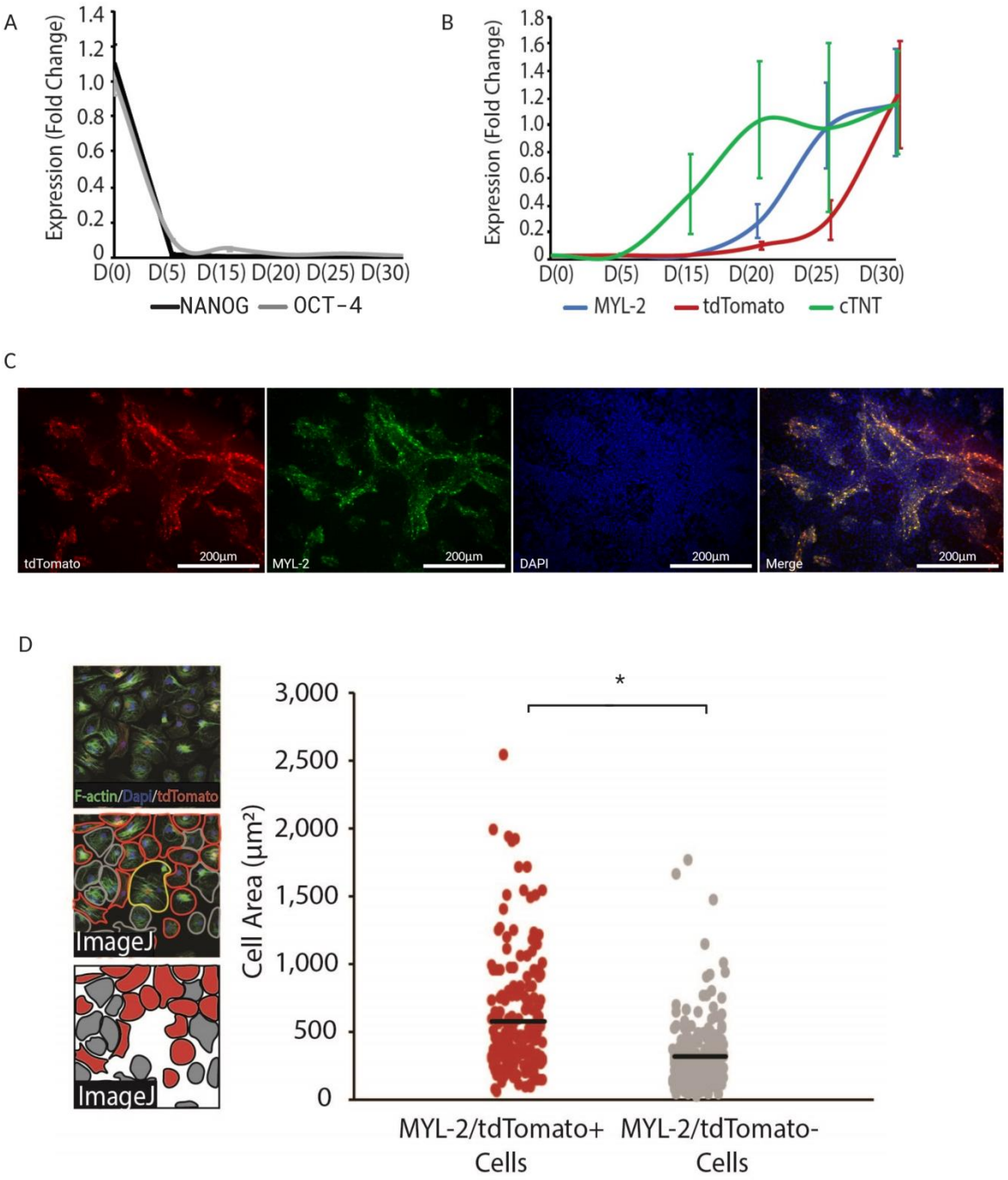


Figure 12: Characterization of the tdTomato positive reporter line. (A) Quantitative RT-PCR of the tdTomato positive cell line for the pluripotency markers *NANOG* and *OCT-4* normalized for *beta-ACTIN () and d30 average ΔCT of the respective gene (ΔΔCT)* (n=3). **(B)** Quantitative RT-PCR of the tdTomato positive cell line for the cardiac markers *cTNT* and *MYL-2*, as well as the fluorophore *tdTomato* normalized for *beta-ACTIN and d30 average ΔCT of the respective gene (ΔΔCT)* (n=3). **(C)** Immunostaining of d30 cardiomyocytes derived from the tdTomato positive cell line for MYL-2, tdTomato fluorophore, DAPI stain for nucleus visualization, and merged image. **(D)** Cell size measurement of tdTomato positive and negative cells after beta-ACTIN immunostaining using the ImageJ software (n=100). *p<0.05

4.5 Functional analysis of the tdTomato positive reporter line

To prove that the *tdTomato* expressing cells behave like ventricular cardiomyocytes electrophysiological analysis was performed. Studies show that the three subtypes of CMs (ventricular, atrial and nodal) generate different action potentials that allow to distinguish them from each other (Zhang, Wilson et al. 2009, Zwi, Caspi et al. 2009, Ma, Guo et al. 2011). Ventricular cardiomyocytes have a more negative maximum diastolic potential (MDP), a rapid action potential upstroke and a distinct plateau phase. Atrial cardiomyocytes can be differentiated from ventricular CMs by the absence of a distinct plateau during repolarization and a higher spontaneous activity frequency. Finally, nodal cardiomyocytes can be distinguished by MDPs that are less negative than those of ventricular and atrial CMs, smaller amplitude action potentials, a slower action potential upstroke, and a pronounced phase-4 depolarization, preceding the action potential upstroke. As shown in figure 13B and C the tdTomato positive d30 cardiomyocytes exhibit electrophysiological characteristics typical for ventricular cardiomyocytes.

Moreover, traction force microscopy was performed in order to obtain results on the contractile force generation of tdTomato positive cells compared to tdTomato negative cells (Fig 13A). By seeding single cardiomyocytes on a specific extracellular matrix (polyacrylamide hydrogels), the displacement of that ECM can be monitored with each contraction, which allows calculating the generated force (Ribeiro, Schwab et al. 2017). TdTomato positive cells generate a significantly higher contractile force compared to tdTomato negative cells.

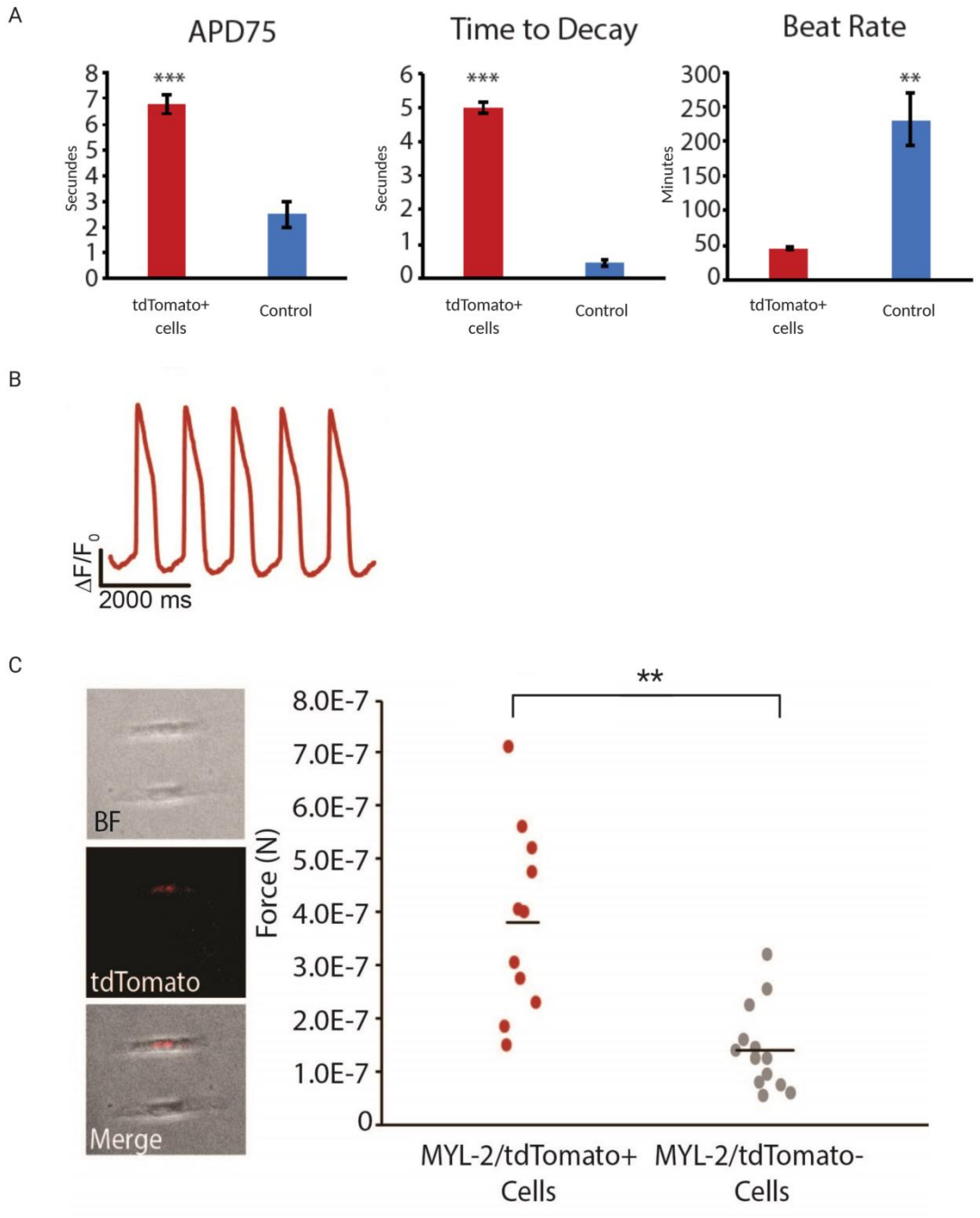


Figure 13: Functional analysis of the tdTomato positive reporter line. (A) Single cell force generation measured by traction force microscopy (n=12). Images of micropatterned cells. **(B)** Representative action potential trace. **(C)** Electrophysiological analysis of unsorted and tdTomato positive cells. **p<0.005; ***p<0.0005

5. Discussion

Recent developments in cardiomyocyte differentiation have optimized and simplified the protocols allowing increased cardiomyocyte yield (Lian, Zhang et al. 2013, Burridge, Matsa et al. 2014). Nevertheless, a truly pure population of CMs, or subtypes of CMs cannot be achieved by sole modifications of differentiation protocols, so far. The widespread cardiomyocyte differentiation protocols are based on the modulation of the Wnt signaling pathway, resulting in 85 to 95% of cardiomyocytes, that can be subdivided in about 60% ventricular, 20% atrial, and 20% nodal CMs (Burridge, Matsa et al. 2014). These numbers are far from a pure population of a cardiomyocyte subtype. Devalla *et al.* used retinoic acid (RA) to enhance the yield of atrial cardiomyocytes, and they were able to achieve a population of up to 85% atrial-like CMs by adding a high concentration (1 µmol/L) of RA between day 4 and 7 of differentiation (Devalla, Schwach et al. 2015). Another study showed, that 77.8% ± 1.7% of non-RA-treated iPSCs were ventricular-like CMs, tantamount to a MYL-2 positive and myosin light chain 2 (MLC2A) negative population. In addition, 10.9% ± 1.6% of the CMs displayed an immature ventricular phenotype being MYL-2 and MLC2A positive (Cyganek, Tiburcy et al. 2018). Although, these studies show that the heterogeneous population after differentiation can be modified, it is clear, that further adjustments have to be made in order to achieve a purified subtype population.

The present work demonstrates that the specific labelling of *MYL-2* with a fluorescent marker (tdTomato) generates a reporter line that allows the unambiguous identification of human iPSC-derived ventricular cardiomyocytes. Nowadays, different methods for creating reporter lines are available. In this study, the recently developed and rapidly expanding genome editing method CRISPR/Cas9 was utilized to label *MYL2* with the fluorophore tdTomato. CRSIPR/Cas9 was developed based on the natural bacterial response to viral infection (Bhaya, Davison et al. 2011, Terns and Terns 2011, Gasiunas, Barrangou et al. 2012, Wiedenheft, Sternberg et al. 2012). This inherent immunological defense mechanism was adapted to allow precise alterations in the genomic sequence of eukaryotic cells. In general, the CRISPR/Cas9 technology allows the induction of double strand breaks (DSB) at almost any desired genomic location. The existence of a DSB within the genome leads to the activation of the cellular endogenous repair machinery - nonhomologous end-joining (NHEJ) and homologous directed repair (HDR) (Syed and Tainer 2018, Wright, Shah et al. 2018,

Marini, Rawal et al. 2019, Carusillo and Mussolino 2020). NHEJ is a mechanism that leads to a random merge of the strands at the site of the DSB resulting in insertions and deletions of nucleotides (indel mutations). Eventually, this may lead to a dysfunctional protein because of induced frameshift mutations or premature stop codons. A more precise repair mechanism is homologous directed repair. HDR ensures a precise repair by using the second, intact allele as repair template. In general, NHEJ occurs much more frequently after DNA damage compared to HDR. In the present study, HDR was used to introduce the desired sequence into the specific genomic location by offering a tailored donor sequence as repair template. Due to rare occurrence of HDR an antibiotic selection cassette (Puromycin) was included into the designed repair template which enables the selection of positive clones that integrated the donor template into their genome. The high percentage of cells with a precise integration of the repair template (78% Fig. 9C) strongly supports the selected method.

It should also be briefly mentioned, that before the rise of the CRISPR/Cas9 technology, other methods for genetic modulations have been used. One widespread method is transcription activator-like effector nucleases (TALENs) (Gaj, Gersbach et al. 2013, Kim and Kim 2014). This method is based on the *FokI* nuclease that is combined with a modifiable DNA binding domain that can cause a nick at a desired location. In order to create a double strand break two TALENs have to be combined. Similar to the CRISPR/Cas9 technology, the nuclease can be navigated to almost any desired genomic location. As described above, this triggers the cellular repair mechanism, either with NHEJ or HDR. Several groups have used the TALEN technology to create reporter lines (Karakikes, Termglinchan et al. 2017, Engels, Olmer et al. 2019, Olmer, Dahlmann et al. 2019). Initially, the TALEN technology was believed to have a lower off-target rate compared to CRISPR/Cas9. Off-target effects are caused by unprecise binding and can lead to unwanted disruption of other genes. There are several studies showing a high-rate of off-target effects of CRISPR/Cas9 (Cradick, Fine et al. 2013, Fu, Foden et al. 2013, Frock, Hu et al. 2015), compared to a low off-target rates using TALENs (Hockemeyer, Wang et al. 2011, Mussolino, Morbitzer et al. 2011). However, one should take into consideration that these initial studies were performed using cancer cell lines where DNA repair mechanisms are broken. Studies performed on iPSCs comparing the off-target rate of CRISPR/Cas9 and TALEN technology with whole-genome sequencing showed a low off-target frequency with both techniques (Smith, Gore et al. 2014, Veres,

Gosis et al. 2014). So far, there is no valid method to predict the occurrence of off-target effects, but Lin *et al*. were able to show that one to three mismatches between genomic DNA and sgRNA can lead to off-target cleavage sites causing indel mutations (Lin, Cradick et al. 2014). The sgRNA used in this present work had 16 different locations within the genomic DNA that showed three or less mismatches. There were no off-targets detectable at any of the location in the edited iPS cell clone.

It has to be pointed out that CRISPR/Cas9 is a technology that is characterized by its ease-of-use and the low cost associated generating a new line. Furthermore using an antibiotic selection cassette, as done in the present work, is time-saving and highly efficient. Since the first publications in 2013, this technology has easily found its way in many labs worldwide and has already become the first choice for genome editing (Cong, Ran et al. 2013, Qi, Larson et al. 2013, Ran, Hsu et al. 2013). In 2015, CRSIPR/Cas9 was even selected by *Science* as Breakthrough of the Year (Frishman 2016). Nevertheless, every genome editing technique goes along with the risk of causing a chromosomal damage (Brunet, Simsek et al. 2009, Simsek, Brunet et al. 2011, Piganeau, Ghezraoui et al. 2013, Kosicki, Tomberg et al. 2018, Cullot, Boutin et al. 2019). It is crucial to test the generated lines before further use, because so far, there is no way of predicting the occurrence of chromosomal disruptions. Karyotyping allows the identification of major chromosomal changes and should, therefore, routinely be done. Furthermore, the sequencing of the edited gene area is indispensable to exclude mistakes during genome editing. The karyogram of tdTomato positive lines generated in the present work showed no chromosomal aberrations after genome editing and sequencing of the edited section excluded mismatches or unintended variants. Furthermore, no alterations of the pluripotency state of the edited iPSCs were detectable.

After making sure the donor sequence was precisely integrated into the genome and that no further damage was caused by genome editing, the generated line was investigated for its ability to differentiate into cardiomyocytes. As described above, a wide range of differentiation protocols are available, neither of them resulting in a pure populations of CMs or subtype of CMs. The initial differentiation protocols relied on the experience with the differentiation of embryonic stem cells and were based on the formation of embryoid bodies (Zhang, Wilson et al. 2009). Unfortunately, these protocols resulted in a very low yield of cardiomyocytes. Based on knowledge from developmental biology regarding

cardiogenesis and the regulation of transcription factors and growth factors, new protocols based on monolayer culture were developed. The differentiation protocol used in this study is based on the modulation of the Wnt signaling pathway by the pathway activator CHIR99021 and the pathway inhibitor Wnt-C59, which results in a cardiomyocyte yield of more than 80%. The CM yield can be further enhanced by glucose starvation based on the findings by Tohyama *et al*., who showed that iPSCs rely on glycolysis for ATP generation (Tohyama, Hattori et al. 2013). Cardiomyocytes, on the other hand, are also able to metabolize pyruvate for energy production. By depleting glucose from the media non-cardiac cells die off, and the CM yield may rise up to 90%.

After differentiation of the edited human iPSC line, flow cytometry analysis revealed that about 55% of the cTNT positive cells at d30 of differentiation were also tdTomato positive. These numbers match the previously described subtype distribution of this differentiation protocol: 60% ventricular, 20% atrial, and 20% nodal CMs (Lian, Zhang et al. 2013, Burridge, Matsa et al. 2014). Furthermore, immunocytochemistry revealed a very convincing overlap of concomitant tdTomato expression and MYL-2 expression, confirming the simultaneous induction of these two genes regulated by the *MYL-2* promoter. As expected, qRT-PCR showed that during differentiation the expression of pluripotency markers was gradually switched off, while the expression of cardiac markers was progressively turned on. These results demonstrate that genome editing did not alter the potential of iPSC to differentiate into cardiomyocytes, and they indicate the labeling of ventricular cardiomyocytes. In order to prove that the tdTomato positive cells are truly ventricular CMs, functional analyses were performed.

In the present work, a new method to evaluate the electrophysiological properties of cardiomyocytes was established. So far, the gold standard for electrophysiological analysis of CMs is patch-clamp analysis (Kornreich 2007, Yechikov, Copaciu et al. 2016, Casini, Verkerk et al. 2017). Patch-clamp analysis is performed by introducing a glass pipette against the cell membrane of an isolated cardiomyocyte. The class pipette is connected to a recording electrode and allows the measurement of the action potential. This method is very complex and has a very steep learning curve. Obtaining electrophysiological properties by using voltage-sensitive dyes on the other hand is a progressive new approach that is characterized by an easier application. This method also allows the evaluation of the action potential of

the cells, having the disadvantage that certain data about ion current characteristics are not able to be obtained, in contrast to patch-clamp analysis. So, the choice of method is based on the question asked. In this study, the question was to evaluate, whether genome edited tdTomato positive cells display action potentials typical of ventricular cardiomyocytes. Electrophysiological analysis using a voltage-sensitive dye is a proven method to answer this question (Casini, Verkerk et al. 2017), and therefore a valid technology to be used in this setting. The data obtained clearly show, that the electrophysiological characteristics of the tdTomato positive cells are indeed those expected for ventricular cardiomyocytes with a distinct plateau phase and accelerated repolarization (Fig.13A and B).

The present work showed that hiPSC-derived ventricular CMs compared to human iPSC-derived non-ventricular CMs are bigger in size. This correlates to existing knowledge about adult atrial and ventricular CMs (Kühnel 2008), emphasizing the similarities of human iPSC-derived CMs and adult mature CMs. Similarly, this study showed, using traction force microscopy, that human iPSC-derived CMs generate a higher force compared to human iPSC-derived non-ventricular CMs. These findings are in good agreement with the characteristics previously described for adult CMs (Wolters 2015).

Taken together, these data show that the tdTomato positive cells display transcriptional and functional characteristics of ventricular cardiomyocytes. Similarly, Zhang *et al.* generated a *Myl-2*-tdTomato knock-in reporter mouse line by electroporating a construct containing IRES-tdTomato-FRT-Neo-FRT flanked by homology arms (Zhang and Nam 2018). The group was able to show, that the *Myl-2* expression is limited to the ventricles throughout the development into the adulthood. However, Myl-2 is not the only discussed ventricular marker in the cardiovascular research field. Nelson *et al*. labeled the iroquois-class homeodomain protein (*Irx4*) using tdTomato in order to study ventricular differentiation (Nelson, Lalit et al. 2016). *Irx4* is one of the earliest markers of ventricular-specific progenitor cells, but its expression is not restricted to ventricular CMs (Bruneau, Bao et al. 2000). *Irx4* expression can also be detected in endothelial and smooth muscle cells. As described above, cardiomyocyte differentiation gives rise to a heterogeneous cell population, not only containing CMs, but also endothelial cells or fibroblasts. Therefore, *Irx4* is not the optimal marker to select for ventricular cardiomyocytes. Nonetheless, one option can be to create a double reporter line, labeling the cardiomyocytes with a cTNT and the ventricular subtype

with Irx4. But, *Myl-2* expression on the other hand is solely restricted to ventricular myocardium (Lee, Ross et al. 1992). Zhang *et al*. used an indirect approach to purify ventricular CMs. The researchers labeled two different proteins, T-box transcription factor 5 (TBX5) and homeobox protein NKX2.5 (NKX2.5) with different fluorophores using CRISPR/Cas9 (Zhang, Termglinchan et al. 2019). TBX5 and NKX2.5 are two important cardiac transcription factors and play a crucial role during heart development. By creating the double-reporter line, the group was able to identify four different subpopulations: $TBX5^{+}NKX2.5^{+}$ display characteristics of cells from the first heart field mainly differentiating into ventricular cardiomyocytes. $TBX5^{+}NKX2.5^{-}$ show characteristics of cells from the epicardial lineage, and can differentiate into nodal CMs. $TBX5^{-}NKX2.5^{+}$ exhibit characteristics of cells from the second heart field, and mainly differentiate into atrial CMs. Finally, $TBX5^{-}$ $NKX2.5^{-}$ generate a subpopulation that displays properties of endothelial cells. Although, the double reporter allows the purification of different subpopulations, neither of them give rise to a pure population of one subtype of cardiomyocytes. So far, labelling MYL2 seems to be the most specific tracking method for ventricular cardiomyocytes.

However, several studies have shown that hiPSC-derived cardiomyocytes are immature, and resemble a rather fetal phenotype (Karakikes, Ameen et al. 2015, Denning, Borgdorff et al. 2016, Goversen, van der Heyden et al. 2018). The creation of this ventricular cardiomyocyte reporter line can thrive the progress of CMs maturity for the ventricular subtype. Being able to truly measure the effects of different protocols and modifications on clearly defined ventricular cardiomyocytes allow the exact assessment of their usefulness. But, driving the cardiomyocyte to a more mature state is not the only field this reporter line may be essential for. As described above, mixed populations of cardiomyocytes are challenging in disease modelling, drug testing and regenerative medicine. When looking at the attempts to transplant cardiac patches onto the ventricles of patients, the fear of arrhythmogenic complications due to the existence of nodal CMs are a big challenge (Shiba, Gomibuchi et al. 2016). Purifying ventricular CMs and creating a patch solely existing out of this subtype, might push the field further. The establishment of a genome edited reporter line which allows tracing of ventricular cardiomyocytes is a powerful tool to obtain a pure subpopulation. But, it must be said, that the purification of iPSC-derived cardiomyocytes using FACS remains challenging (Waas, Weerasekera et al. 2019). One general feature of cardiomyocytes is their larger size (100–150 x 20–35 µm) (Vu T.D. 2014), requiring the use of

bigger nozzles during FACS. But, also large nozzles can cause damage or even destruction of cardiomyocytes. One method to maintain cell viability is the use of large particle FACS (LP-FACS) (Lopez, Sharma et al. 2017, Kannan, Miyamoto et al. 2019) using nozzle with 200µm or even 500 µm.

6. Conclusion

The CRISPR/Cas9 technology allows the efficient and reproducible generation of a ventricular cardiomyocyte reporter line. This study was able to demonstrate that the labeled cells display transcriptional and functional characteristics of ventricular cardiomyocytes. Furthermore, there were no chromosomal damage or alterations of the pluripotency state of the edited iPSCs detectable. The generated reporter line is an important instrument in toolbox of cardiovascular research.

7. Bibliography

Anokye-Danso, F., C. M. Trivedi, D. Juhr, M. Gupta, Z. Cui, Y. Tian, Y. Zhang, W. Yang, P. J. Gruber, J. A. Epstein and E. E. Morrisey (2011). "Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency." Cell Stem Cell **8**(4): 376-388.

Barton, P. J. and M. E. Buckingham (1985). "The myosin alkali light chain proteins and their genes." Biochem J **231**(2): 249-261.

Bellin, M., S. Casini, R. P. Davis, C. D'Aniello, J. Haas, D. Ward-van Oostwaard, L. G. Tertoolen, C. B. Jung, D. A. Elliott, A. Welling, K. L. Laugwitz, A. Moretti and C. L. Mummery (2013). "Isogenic human pluripotent stem cell pairs reveal the role of a KCNH2 mutation in long-QT syndrome." EMBO J **32**(24): 3161-3175.

Beltrami, A. P., L. Barlucchi, D. Torella, M. Baker, F. Limana, S. Chimenti, H. Kasahara, M. Rota, E. Musso, K. Urbanek, A. Leri, J. Kajstura, B. Nadal-Ginard and P. Anversa (2003). "Adult cardiac stem cells are multipotent and support myocardial regeneration." Cell **114**(6): 763-776.

Bhaya, D., M. Davison and R. Barrangou (2011). "CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation." Annu Rev Genet **45**: 273-297.

Bruneau, B. G., Z. Z. Bao, M. Tanaka, J. J. Schott, S. Izumo, C. L. Cepko, J. G. Seidman and C. E. Seidman (2000). "Cardiac expression of the ventricle-specific homeobox gene Irx4 is modulated by Nkx2-5 and dHand." Dev Biol **217**(2): 266-277.

Brunet, E., D. Simsek, M. Tomishima, R. DeKelver, V. M. Choi, P. Gregory, F. Urnov, D. M. Weinstock and M. Jasin (2009). "Chromosomal translocations induced at specified loci in human stem cells." Proc Natl Acad Sci U S A **106**(26): 10620-10625.

Burridge, P. W., E. Matsa, P. Shukla, Z. C. Lin, J. M. Churko, A. D. Ebert, F. Lan, S. Diecke, B. Huber, N. M. Mordwinkin, J. R. Plews, O. J. Abilez, B. Cui, J. D. Gold and J. C. Wu (2014). "Chemically defined generation of human cardiomyocytes." Nat Methods **11**(8): 855-860.

Carusillo, A. and C. Mussolino (2020). "DNA Damage: From Threat to Treatment." Cells **9**(7).

Carvajal-Vergara, X., A. Sevilla, S. L. D'Souza, Y. S. Ang, C. Schaniel, D. F. Lee, L. Yang, A. D. Kaplan, E. D. Adler, R. Rozov, Y. Ge, N. Cohen, L. J. Edelmann, B. Chang, A. Waghray, J. Su, S. Pardo, K. D. Lichtenbelt, M. Tartaglia, B. D. Gelb and I. R. Lemischka (2010). "Patient-specific induced pluripotent stem-cell-derived models of LEOPARD syndrome." Nature **465**(7299): 808-812.

Casini, S., A. O. Verkerk and C. A. Remme (2017). "Human iPSC-Derived Cardiomyocytes for Investigation of Disease Mechanisms and Therapeutic Strategies in Inherited Arrhythmia Syndromes: Strengths and Limitations." Cardiovasc Drugs Ther **31**(3): 325-344.

Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini and F. Zhang (2013). "Multiplex genome engineering using CRISPR/Cas systems." Science **339**(6121): 819-823.

Cradick, T. J., E. J. Fine, C. J. Antico and G. Bao (2013). "CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity." Nucleic Acids Res **41**(20): 9584-9592.

Cullot, G., J. Boutin, J. Toutain, F. Prat, P. Pennamen, C. Rooryck, M. Teichmann, E. Rousseau, I. Lamrissi-Garcia, V. Guyonnet-Duperat, A. Bibeyran, M. Lalanne, V. Prouzet-Mauleon, B. Turcq, C. Ged, J. M. Blouin, E. Richard, S. Dabernat, F. Moreau-Gaudry and A. Bedel (2019).

"CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations." Nat Commun **10**(1): 1136.
Cyganek, L., M. Tiburcy, K. Sekeres, K. Gerstenberg, H. Bohnenberger, C. Lenz, S. Henze, M. Stauske, G. Salinas, W. H. Zimmermann, G. Hasenfuss and K. Guan (2018). "Deep phenotyping of human induced pluripotent stem cell-derived atrial and ventricular cardiomyocytes." JCI Insight **3**(12).
Denning, C., V. Borgdorff, J. Crutchley, K. S. Firth, V. George, S. Kalra, A. Kondrashov, M. D. Hoang, D. Mosqueira, A. Patel, L. Prodanov, D. Rajamohan, W. C. Skarnes, J. G. Smith and L. E. Young (2016). "Cardiomyocytes from human pluripotent stem cells: From laboratory curiosity to industrial biomedical platform." Biochim Biophys Acta **1863**(7 Pt B): 1728-1748.
Devalla, H. D. and R. Passier (2018). "Cardiac differentiation of pluripotent stem cells and implications for modeling the heart in health and disease." Sci Transl Med **10**(435).
Devalla, H. D., V. Schwach, J. W. Ford, J. T. Milnes, S. El-Haou, C. Jackson, K. Gkatzis, D. A. Elliott, S. M. Chuva de Sousa Lopes, C. L. Mummery, A. O. Verkerk and R. Passier (2015). "Atrial-like cardiomyocytes from human pluripotent stem cells are a robust preclinical model for assessing atrial-selective pharmacology." EMBO Mol Med **7**(4): 394-410.
Engels, L., R. Olmer, J. de la Roche, G. Gohring, S. Ulrich, R. Haller, U. Martin and S. Merkert (2019). "Generation of a CFTR knock-in reporter cell line (MHHi006-A-1) from a human induced pluripotent stem cell line." Stem Cell Res **40**: 101542.
Frishman, R. (2016). "Cover stories: Making the Breakthrough of the Year cover." Science **354**(6319): 1497.
Frock, R. L., J. Hu, R. M. Meyers, Y. J. Ho, E. Kii and F. W. Alt (2015). "Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases." Nat Biotechnol **33**(2): 179-186.
Fu, Y., J. A. Foden, C. Khayter, M. L. Maeder, D. Reyon, J. K. Joung and J. D. Sander (2013). "High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells." Nat Biotechnol **31**(9): 822-826.
Fusaki, N., H. Ban, A. Nishiyama, K. Saeki and M. Hasegawa (2009). "Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome." Proc Jpn Acad Ser B Phys Biol Sci **85**(8): 348-362.
Gaj, T., C. A. Gersbach and C. F. Barbas, 3rd (2013). "ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering." Trends Biotechnol **31**(7): 397-405.
Gasiunas, G., R. Barrangou, P. Horvath and V. Siksnys (2012). "Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria." Proc Natl Acad Sci U S A **109**(39): E2579-2586.
Gawlik-Rzemieniewska, N. and I. Bednarek (2016). "The role of NANOG transcriptional factor in the development of malignant phenotype of cancer cells." Cancer Biol Ther **17**(1): 1-10.
Goversen, B., M. A. G. van der Heyden, T. A. B. van Veen and T. P. de Boer (2018). "The immature electrophysiological phenotype of iPSC-CMs still hampers in vitro drug screening: Special focus on IK1." Pharmacol Ther **183**: 127-136.
Gupta, R. M., T. B. Meissner, C. A. Cowan and K. Musunuru (2016). "Genome-Edited Human Pluripotent Stem Cell-Derived Macrophages as a Model of Reverse Cholesterol Transport--Brief Report." Arterioscler Thromb Vasc Biol **36**(1): 15-18.
Harvey, W. (1628). Exercitatio Anatomica de Motu Cordis et Sanguinis in Animalibus.
Heath, E. I., L. K. Heilbrun, D. Smith, W. M. Schopperle, Y. Ju, S. Bolton, Q. Ahmed and W. A. Sakr (2018). "Overexpression of the Pluripotent Stem Cell Marker Podocalyxin in Prostate Cancer." Anticancer Res **38**(11): 6361-6366.

Hinson, J. T., A. Chopra, N. Nafissi, W. J. Polacheck, C. C. Benson, S. Swist, J. Gorham, L. Yang, S. Schafer, C. C. Sheng, A. Haghighi, J. Homsy, N. Hubner, G. Church, S. A. Cook, W. A. Linke, C. S. Chen, J. G. Seidman and C. E. Seidman (2015). "HEART DISEASE. Titin mutations in iPS cells define sarcomere insufficiency as a cause of dilated cardiomyopathy." Science **349**(6251): 982-986.
Hockemeyer, D., H. Wang, S. Kiani, C. S. Lai, Q. Gao, J. P. Cassady, G. J. Cost, L. Zhang, Y. Santiago, J. C. Miller, B. Zeitler, J. M. Cherone, X. Meng, S. J. Hinkley, E. J. Rebar, P. D. Gregory, F. D. Urnov and R. Jaenisch (2011). "Genetic engineering of human pluripotent cells using TALE nucleases." Nat Biotechnol **29**(8): 731-734.
Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna and E. Charpentier (2012). "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Science **337**(6096): 816-821.
Kannan, S., M. Miyamoto, B. L. Lin, R. Zhu, S. Murphy, D. A. Kass, P. Andersen and C. Kwon (2019). "Large Particle Fluorescence-Activated Cell Sorting Enables High-Quality Single-Cell RNA Sequencing and Functional Analysis of Adult Cardiomyocytes." Circ Res **125**(5): 567-569.
Karakikes, I., M. Ameen, V. Termglinchan and J. C. Wu (2015). "Human induced pluripotent stem cell-derived cardiomyocytes: insights into molecular, cellular, and functional phenotypes." Circ Res **117**(1): 80-88.
Karakikes, I., F. Stillitano, M. Nonnenmacher, C. Tzimas, D. Sanoudou, V. Termglinchan, C. W. Kong, S. Rushing, J. Hansen, D. Ceholski, F. Kolokathis, D. Kremastinos, A. Katoulis, L. Ren, N. Cohen, J. Gho, D. Tsiapras, A. Vink, J. C. Wu, F. W. Asselbergs, R. A. Li, J. S. Hulot, E. G. Kranias and R. J. Hajjar (2015). "Correction of human phospholamban R14del mutation associated with cardiomyopathy using targeted nucleases and combination therapy." Nat Commun **6**: 6955.
Karakikes, I., V. Termglinchan, D. A. Cepeda, J. Lee, S. Diecke, A. Hendel, I. Itzhaki, M. Ameen, R. Shrestha, H. Wu, N. Ma, N. Y. Shao, T. Seeger, N. Woo, K. D. Wilson, E. Matsa, M. H. Porteus, V. Sebastiano and J. C. Wu (2017). "A Comprehensive TALEN-Based Knockout Library for Generating Human-Induced Pluripotent Stem Cell-Based Models for Cardiovascular Diseases." Circ Res **120**(10): 1561-1571.
Kattman, S. J., A. D. Witty, M. Gagliardi, N. C. Dubois, M. Niapour, A. Hotta, J. Ellis and G. Keller (2011). "Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines." Cell Stem Cell **8**(2): 228-240.
Kim, H. and J. S. Kim (2014). "A guide to genome engineering with programmable nucleases." Nat Rev Genet **15**(5): 321-334.
Kornreich, B. G. (2007). "The patch clamp technique: principles and technical considerations." J Vet Cardiol **9**(1): 25-37.
Kosicki, M., K. Tomberg and A. Bradley (2018). "Erratum: Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements." Nat Biotechnol **36**(9): 899.
Kühnel, W. (2008). Taschenatlas Histologie, Kühnel W.
Lee, G. and I. Saito (1998). "Role of nucleotide sequences of loxP spacer region in Cre-mediated recombination." Gene **216**(1): 55-65.
Lee, K. J., R. S. Ross, H. A. Rockman, A. N. Harris, T. X. O'Brien, M. van Bilsen, H. E. Shubeita, R. Kandolf, G. Brem, J. Price and et al. (1992). "Myosin light chain-2 luciferase transgenic mice reveal distinct regulatory programs for cardiac and skeletal muscle-specific expression of a single contractile protein gene." J Biol Chem **267**(22): 15875-15885.

Lian, X., C. Hsiao, G. Wilson, K. Zhu, L. B. Hazeltine, S. M. Azarin, K. K. Raval, J. Zhang, T. J. Kamp and S. P. Palecek (2012). "Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling." Proc Natl Acad Sci U S A **109**(27): E1848-1857.
Lian, X., J. Zhang, S. M. Azarin, K. Zhu, L. B. Hazeltine, X. Bao, C. Hsiao, T. J. Kamp and S. P. Palecek (2013). "Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/beta-catenin signaling under fully defined conditions." Nat Protoc **8**(1): 162-175.
Lin, H., Q. Du, Q. Li, O. Wang, Z. Wang, C. Elowsky, K. Liu, C. Zhang, S. Chung, B. Duan and Y. Lei (2018). "Manufacturing human pluripotent stem cell derived endothelial cells in scalable and cell-friendly microenvironments." Biomater Sci **7**(1): 373-388.
Lin, Y., T. J. Cradick, M. T. Brown, H. Deshmukh, P. Ranjan, N. Sarode, B. M. Wile, P. M. Vertino, F. J. Stewart and G. Bao (2014). "CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences." Nucleic Acids Res **42**(11): 7473-7485.
Lopez, J. E., J. Sharma, J. Avila, T. S. Wood, J. E. VanDyke, B. McLaughlin, C. K. Abbey, A. Wong, B. E. Myagmar, P. M. Swigart, P. C. Simpson and N. Chiamvimonvat (2017). "Novel large-particle FACS purification of adult ventricular myocytes reveals accumulation of myosin and actin disproportionate to cell size and proteome in normal post-weaning development." J Mol Cell Cardiol **111**: 114-122.
Ma, J., L. Guo, S. J. Fiene, B. D. Anson, J. A. Thomson, T. J. Kamp, K. L. Kolaja, B. J. Swanson and C. T. January (2011). "High purity human-induced pluripotent stem cell-derived cardiomyocytes: electrophysiological properties of action potentials and ionic currents." Am J Physiol Heart Circ Physiol **301**(5): H2006-2017.
Marini, F., C. C. Rawal, G. Liberi and A. Pellicioli (2019). "Regulation of DNA Double Strand Breaks Processing: Focus on Barriers." Front Mol Biosci **6**: 55.
McKeithan, W. L., A. Savchenko, M. S. Yu, F. Cerignoli, A. A. N. Bruyneel, J. H. Price, A. R. Colas, E. W. Miller, J. R. Cashman and M. Mercola (2017). "An Automated Platform for Assessment of Congenital and Drug-Induced Arrhythmia with hiPSC-Derived Cardiomyocytes." Front Physiol **8**: 766.
Moretti, A., M. Bellin, A. Welling, C. B. Jung, J. T. Lam, L. Bott-Flugel, T. Dorn, A. Goedel, C. Hohnke, F. Hofmann, M. Seyfarth, D. Sinnecker, A. Schomig and K. L. Laugwitz (2010). "Patient-specific induced pluripotent stem-cell models for long-QT syndrome." N Engl J Med **363**(15): 1397-1409.
Mussolino, C., R. Morbitzer, F. Lutge, N. Dannemann, T. Lahaye and T. Cathomen (2011). "A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity." Nucleic Acids Res **39**(21): 9283-9293.
Musunuru, K., F. Sheikh, R. M. Gupta, S. R. Houser, K. O. Maher, D. J. Milan, A. Terzic, J. C. Wu, G. American Heart Association Council on Functional, B. Translational, Y. Council on Cardiovascular Disease in the, C. Council on and N. Stroke (2018). "Induced Pluripotent Stem Cells for Cardiovascular Disease Modeling and Precision Medicine: A Scientific Statement From the American Heart Association." Circ Genom Precis Med **11**(1): e000043.
Nelson, D. O., P. A. Lalit, M. Biermann, Y. S. Markandeya, D. L. Capes, L. Addesso, G. Patel, T. Han, M. C. John, P. A. Powers, K. M. Downs, T. J. Kamp and G. E. Lyons (2016). "Irx4 Marks a Multipotent, Ventricular-Specific Progenitor Cell." Stem Cells **34**(12): 2875-2888.
Okano, S. and Y. Shiba (2019). "Therapeutic Potential of Pluripotent Stem Cells for Cardiac Repair after Myocardial Infarction." Biol Pharm Bull **42**(4): 524-530.

Okita, K., M. Nakagawa, H. Hyenjong, T. Ichisaka and S. Yamanaka (2008). "Generation of mouse induced pluripotent stem cells without viral vectors." Science **322**(5903): 949-953.
Olmer, R., J. Dahlmann, S. Merkert, S. Baus, G. Gohring and U. Martin (2019). "Generation of a NKX2.1 knock-in reporter cell line from human induced pluripotent stem cells (MHHi006-A-2)." Stem Cell Res **39**: 101492.
Orban, P. C., D. Chui and J. D. Marth (1992). "Tissue- and site-specific DNA recombination in transgenic mice." Proc Natl Acad Sci U S A **89**(15): 6861-6865.
Pfaffl, M. (2004). Quantification strategies in real-time PCR. A-Z of quantitative PCR. S. A. Bustin. La Jolla, CA, USA**:** 87-112.
Piganeau, M., H. Ghezraoui, A. De Cian, L. Guittat, M. Tomishima, L. Perrouault, O. Rene, G. E. Katibah, L. Zhang, M. C. Holmes, Y. Doyon, J. P. Concordet, C. Giovannangeli, M. Jasin and E. Brunet (2013). "Cancer translocations in human cells induced by zinc finger and TALE nucleases." Genome Res **23**(7): 1182-1193.
Protze, S. I., J. Liu, U. Nussinovitch, L. Ohana, P. H. Backx, L. Gepstein and G. M. Keller (2017). "Sinoatrial node cardiomyocytes derived from human pluripotent cells function as a biological pacemaker." Nat Biotechnol **35**(1): 56-68.
Qi, L. S., M. H. Larson, L. A. Gilbert, J. A. Doudna, J. S. Weissman, A. P. Arkin and W. A. Lim (2013). "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression." Cell **152**(5): 1173-1183.
Ran, F. A., P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott and F. Zhang (2013). "Genome engineering using the CRISPR-Cas9 system." Nat Protoc **8**(11): 2281-2308.
Ribeiro, A. J. S., O. Schwab, M. A. Mandegar, Y. S. Ang, B. R. Conklin, D. Srivastava and B. L. Pruitt (2017). "Multi-Imaging Method to Assay the Contractile Mechanical Output of Micropatterned Human iPSC-Derived Cardiac Myocytes." Circ Res **120**(10): 1572-1583.
Schopperle, W. M. and W. C. DeWolf (2007). "The TRA-1-60 and TRA-1-81 human pluripotent stem cell markers are expressed on podocalyxin in embryonal carcinoma." Stem Cells **25**(3): 723-730.
Scuderi, G. J. and J. Butcher (2017). "Naturally Engineered Maturation of Cardiomyocytes." Front Cell Dev Biol **5**: 50.
Sheikh, F., R. C. Lyon and J. Chen (2015). "Functions of myosin light chain-2 (MYL2) in cardiac muscle and disease." Gene **569**(1): 14-20.
Sheikh, F., K. Ouyang, S. G. Campbell, R. C. Lyon, J. Chuang, D. Fitzsimons, J. Tangney, C. G. Hidalgo, C. S. Chung, H. Cheng, N. D. Dalton, Y. Gu, H. Kasahara, M. Ghassemian, J. H. Omens, K. L. Peterson, H. L. Granzier, R. L. Moss, A. D. McCulloch and J. Chen (2012). "Mouse and computational models link Mlc2v dephosphorylation to altered myosin kinetics in early cardiac disease." J Clin Invest **122**(4): 1209-1221.
Shiba, Y., T. Gomibuchi, T. Seto, Y. Wada, H. Ichimura, Y. Tanaka, T. Ogasawara, K. Okada, N. Shiba, K. Sakamoto, D. Ido, T. Shiina, M. Ohkura, J. Nakai, N. Uno, Y. Kazuki, M. Oshimura, I. Minami and U. Ikeda (2016). "Allogeneic transplantation of iPS cell-derived cardiomyocytes regenerates primate hearts." Nature **538**(7625): 388-391.
Simsek, D., E. Brunet, S. Y. Wong, S. Katyal, Y. Gao, P. J. McKinnon, J. Lou, L. Zhang, J. Li, E. J. Rebar, P. D. Gregory, M. C. Holmes and M. Jasin (2011). "DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation." PLoS Genet **7**(6): e1002080.
Smith, C., A. Gore, W. Yan, L. Abalde-Atristain, Z. Li, C. He, Y. Wang, R. A. Brodsky, K. Zhang, L. Cheng and Z. Ye (2014). "Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs." Cell Stem Cell **15**(1): 12-13.

Stadtfeld, M., M. Nagaya, J. Utikal, G. Weir and K. Hochedlinger (2008). "Induced pluripotent stem cells generated without viral integration." Science **322**(5903): 945-949.
Sun, N., M. Yazawa, J. Liu, L. Han, V. Sanchez-Freire, O. J. Abilez, E. G. Navarrete, S. Hu, L. Wang, A. Lee, A. Pavlovic, S. Lin, R. Chen, R. J. Hajjar, M. P. Snyder, R. E. Dolmetsch, M. J. Butte, E. A. Ashley, M. T. Longaker, R. C. Robbins and J. C. Wu (2012). "Patient-specific induced pluripotent stem cells as a model for familial dilated cardiomyopathy." Sci Transl Med **4**(130): 130ra147.
Syed, A. and J. A. Tainer (2018). "The MRE11-RAD50-NBS1 Complex Conducts the Orchestration of Damage Signaling and Outcomes to Stress in DNA Replication and Repair." Annu Rev Biochem **87**: 263-294.
Tai, M. H., C. C. Chang, M. Kiupel, J. D. Webster, L. K. Olson and J. E. Trosko (2005). "Oct4 expression in adult human stem cells: evidence in support of the stem cell theory of carcinogenesis." Carcinogenesis **26**(2): 495-502.
Takahashi, K., K. Tanabe, M. Ohnuki, M. Narita, T. Ichisaka, K. Tomoda and S. Yamanaka (2007). "Induction of pluripotent stem cells from adult human fibroblasts by defined factors." Cell **131**(5): 861-872.
Terns, M. P. and R. M. Terns (2011). "CRISPR-based adaptive immune systems." Curr Opin Microbiol **14**(3): 321-327.
Theodoris, C. V., M. Li, M. P. White, L. Liu, D. He, K. S. Pollard, B. G. Bruneau and D. Srivastava (2015). "Human disease modeling reveals integrated transcriptional and epigenetic mechanisms of NOTCH1 haploinsufficiency." Cell **160**(6): 1072-1086.
Tohyama, S., F. Hattori, M. Sano, T. Hishiki, Y. Nagahata, T. Matsuura, H. Hashimoto, T. Suzuki, H. Yamashita, Y. Satoh, T. Egashira, T. Seki, N. Muraoka, H. Yamakawa, Y. Ohgino, T. Tanaka, M. Yoichi, S. Yuasa, M. Murata, M. Suematsu and K. Fukuda (2013). "Distinct metabolic flow enables large-scale purification of mouse and human pluripotent stem cell-derived cardiomyocytes." Cell Stem Cell **12**(1): 127-137.
Veres, A., B. S. Gosis, Q. Ding, R. Collins, A. Ragavendran, H. Brand, S. Erdin, C. A. Cowan, M. E. Talkowski and K. Musunuru (2014). "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 **15**(1): 27-30.
Vu T.D. , K. T. (2014). Cardiac Regeneration and Repair.
Waas, M., R. Weerasekera, E. M. Kropp, M. Romero-Tejeda, E. N. Poon, K. R. Boheler, P. W. Burridge and R. L. Gundry (2019). "Are These Cardiomyocytes? Protocol Development Reveals Impact of Sample Preparation on the Accuracy of Identifying Cardiomyocytes by Flow Cytometry." Stem Cell Reports **12**(2): 395-410.
Waldron, D. (2017). "Gene therapy: In vivo gene editing in non-dividing cells." Nat Rev Genet **18**(1): 1.
Wang, G., M. L. McCain, L. Yang, A. He, F. S. Pasqualini, A. Agarwal, H. Yuan, D. Jiang, D. Zhang, L. Zangi, J. Geva, A. E. Roberts, Q. Ma, J. Ding, J. Chen, D. Z. Wang, K. Li, J. Wang, R. J. Wanders, W. Kulik, F. M. Vaz, M. A. Laflamme, C. E. Murry, K. R. Chien, R. I. Kelley, G. M. Church, K. K. Parker and W. T. Pu (2014). "Modeling the mitochondrial cardiomyopathy of Barth syndrome with induced pluripotent stem cell and heart-on-chip technologies." Nat Med **20**(6): 616-623.
Wang, Y., F. Wang, R. Wang, P. Zhao and Q. Xia (2015). "2A self-cleaving peptide-based multi-gene expression system in the silkworm Bombyx mori." Sci Rep **5**: 16273.
Wiedenheft, B., S. H. Sternberg and J. A. Doudna (2012). "RNA-guided genetic silencing systems in bacteria and archaea." Nature **482**(7385): 331-338.

Wolters, B. (2015). Zell-biomechanische Untersuchungen von Modellen des Herzmuskelgewebes. Dr. rer. nat., Rheinischen Friedrich-Wilhelms-Universität Bonn.
World Heart Organization. (2017). "Cardiovascular diseases (CVDs)." from https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).
Wright, W. D., S. S. Shah and W. D. Heyer (2018). "Homologous recombination and the repair of DNA double-strand breaks." J Biol Chem **293**(27): 10524-10535.
Yazawa, M., B. Hsueh, X. Jia, A. M. Pasca, J. A. Bernstein, J. Hallmayer and R. E. Dolmetsch (2011). "Using induced pluripotent stem cells to investigate cardiac phenotypes in Timothy syndrome." Nature **471**(7337): 230-234.
Yechikov, S., R. Copaciu, J. M. Gluck, W. Deng, N. Chiamvimonvat, J. W. Chan and D. K. Lieu (2016). "Same-Single-Cell Analysis of Pacemaker-Specific Markers in Human Induced Pluripotent Stem Cell-Derived Cardiomyocyte Subtypes Classified by Electrophysiology." Stem Cells **34**(11): 2670-2680.
Zhang, J., G. F. Wilson, A. G. Soerens, C. H. Koonce, J. Yu, S. P. Palecek, J. A. Thomson and T. J. Kamp (2009). "Functional cardiomyocytes derived from human induced pluripotent stem cells." Circ Res **104**(4): e30-41.
Zhang, J. Z., V. Termglinchan, N. Y. Shao, I. Itzhaki, C. Liu, N. Ma, L. Tian, V. Y. Wang, A. C. Y. Chang, H. Guo, T. Kitani, H. Wu, C. K. Lam, K. Kodo, N. Sayed, H. M. Blau and J. C. Wu (2019). "A Human iPSC Double-Reporter System Enables Purification of Cardiac Lineage Subpopulations with Distinct Function and Drug Response Profiles." Cell Stem Cell **24**(5): 802-811 e805.
Zhang, Z. and Y. J. Nam (2018). "Generation of MLC-2v-tdTomato knock-in reporter mouse line." Genesis **56**(10): e23256.
Zwi-Dantsis, L. and L. Gepstein (2012). "Induced pluripotent stem cells for cardiac repair." Cell Mol Life Sci **69**(19): 3285-3299.
Zwi, L., O. Caspi, G. Arbel, I. Huber, A. Gepstein, I. H. Park and L. Gepstein (2009). "Cardiomyocyte differentiation of human induced pluripotent stem cells." Circulation **120**(15): 1513-1523.

8. Annex

8.1 Supplement

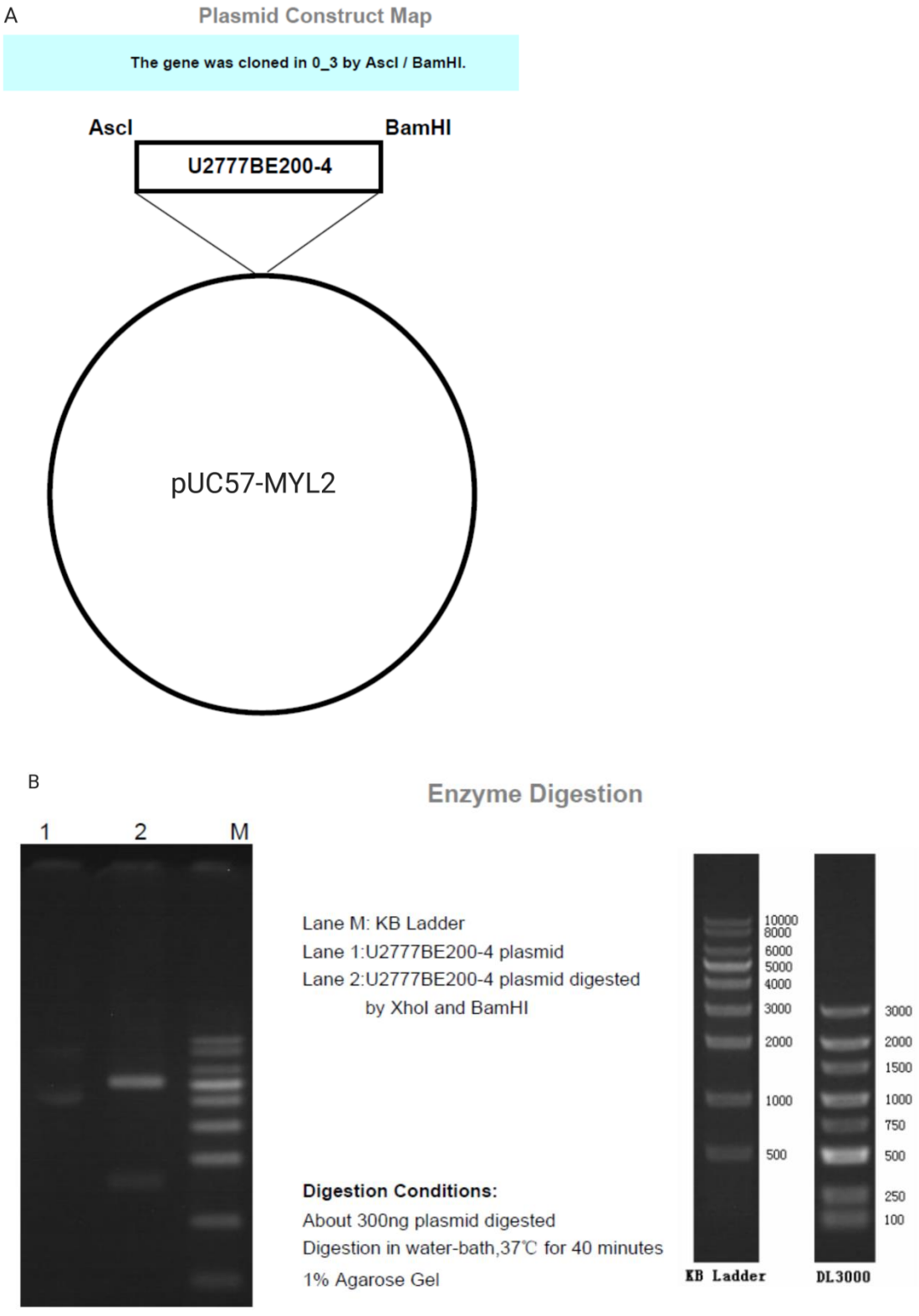


Figure S1: Vector design of the donor construct pUC27-MYL2. (A) Plasmid construct map with the insertion site of the donor sequence **(B)** Enzyme digestion as quality control.

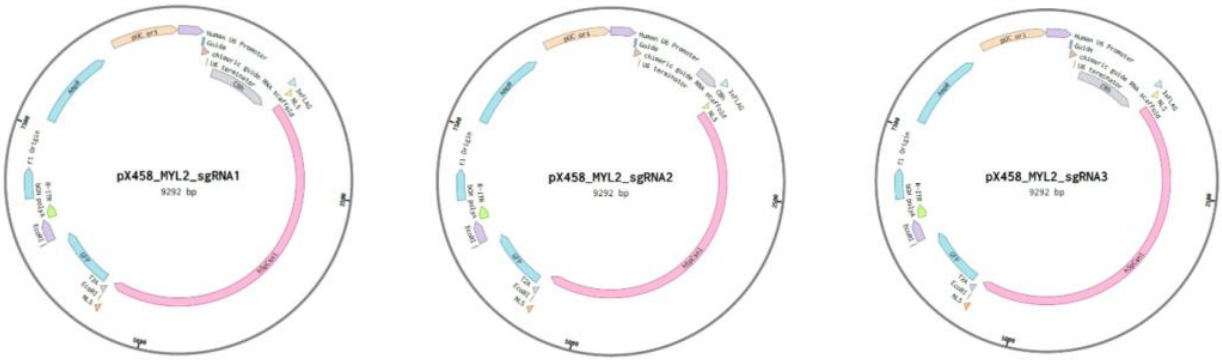


Figure S2: Maps of the pX458 vector with the different sgRNAs.

A)

AAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTAC
CGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGACGGAG
AAGAGAAGGACTAGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG
AAAAAGTGGCACCGAGTCGGTGCTTTTTTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAG
TCCGTTTTTAGCGCGTGCGCCAATTCTGCAGACAAATGGCTCTAGAGGTACCCGTTACATAACTTACG
GTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGT
GTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTGTGCCC
AGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCTCCCCACCCCC

B)

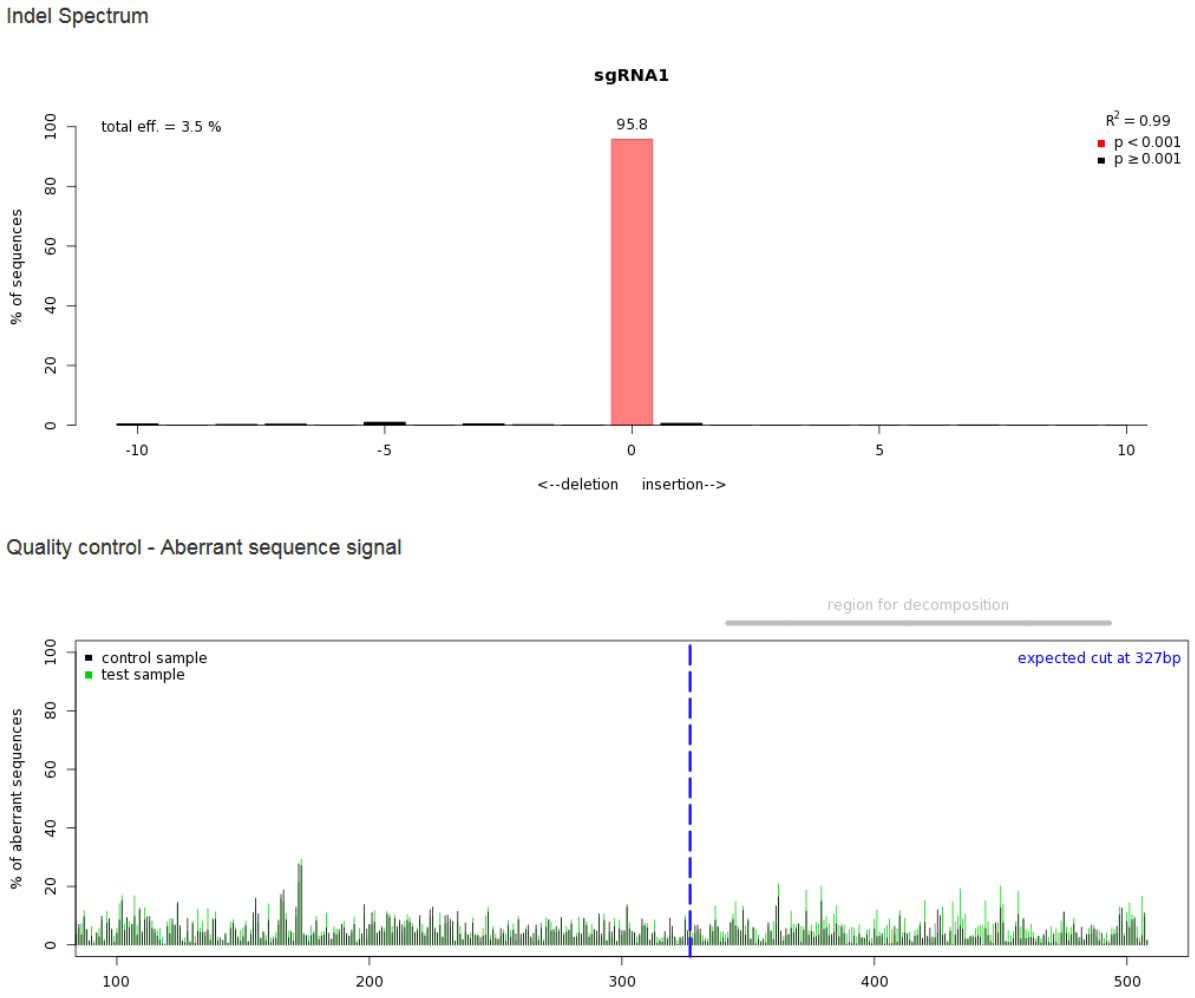
AAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTAC
CGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGCGGAGA
AGAGAAGGACTAGGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG
AAAAAGTGGCACCGAGTCGGTGCTTTTTTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAG
TCCGTTTTTAGCGCGTGCGCCAATTCTGCAGACAAATGGCTCTAGAGGTACCCGTTACATAACTTACG
GTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGT
GTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTGTGCCC
AGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCTCCCCACCCCC

C)

AAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTAC
CGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGAGAAGA
GAAGGACTAGGAGGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG
AAAAAGTGGCACCGAGTCGGTGCTTTTTTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAG
TCCGTTTTTAGCGCGTGCGCCAATTCTGCAGACAAATGGCTCTAGAGGTACCCGTTACATAACTTACG
GTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGT
GTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTGTGCCC

AGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCTCCCCACCCCC

Figure S3: Sequencing data of the inserted sgRNA sequence into the pX458 vector. The red labeled sequence marks the sgRNA. **(A)** pX458_sgRNA1_clone1. **(B)** pX458_sgRNA2_clone1. **(C)** pX458_sgRNA3_clone1.



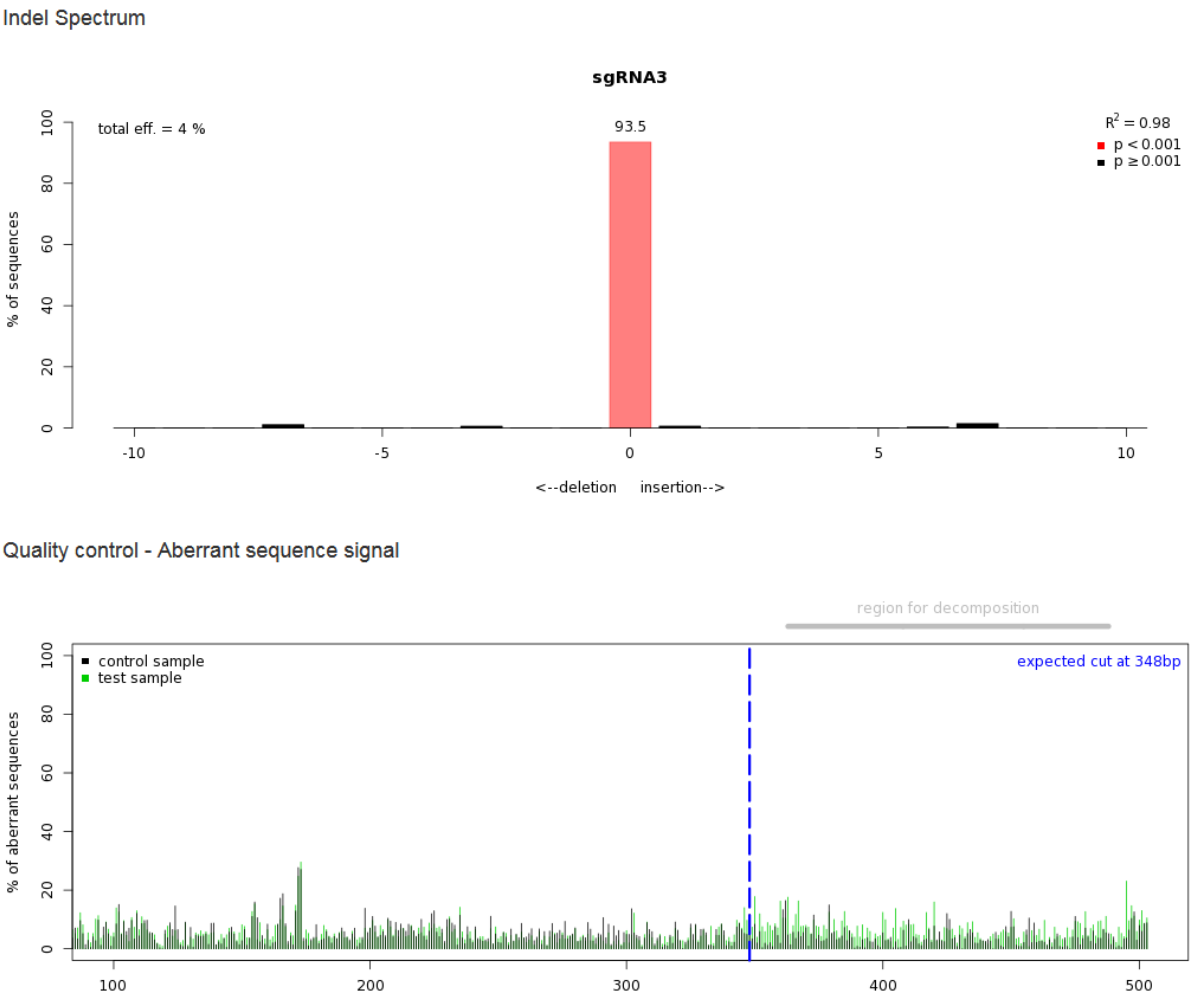


Figure S4: Cutting efficiency of sgRNA1 and 3 calculated using the TIDE software. The upper graphs depict the estimated cutting efficiency with 3.5% for sgRNA1 and 4% for sgRNA3. The lower graphs show that the aberrant sequence signals start at the expected cutting site of the sgRNAs.

A)

Homo sapiens X BAC RP11-109P4 (Roswell Park Cancer Institute Human BAC Library) complete sequence

Sequence ID: AC097261.6 **Length:** 97628 **Number of Matches:** 1

Range 1: 63017 to 63463 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
807 bits(894)	0.0	447/447(100%)	0/447(0%)	Plus/Plus

```
Query  1      CAGATAGGACTGGAAGCAGGACTGGGATGATACAGGGATATGAGCATGTGTCCACCCGTA  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  63017  CAGATAGGACTGGAAGCAGGACTGGGATGATACAGGGATATGAGCATGTGTCCACCCGTA  63076

Query  61     GATCTCTTGTCACTGACTCTACCTAGTCCTTCTCTTCTCCTTAGTTCACTGTCCTCAGTA  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  63077  GATCTCTTGTCACTGACTCTACCTAGTCCTTCTCTTCTCCTTAGTTCACTGTCCTCAGTA  63136

Query  121    AACAGCTTTTCCTGGTCTTCTGTTCACAAGGGATAGCCCTCACTCTGGGAGACTACTCCT  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  63137  AACAGCTTTTCCTGGTCTTCTGTTCACAAGGGATAGCCCTCACTCTGGGAGACTACTCCT  63196

Query  181    CTCCTGGAAATTCCATATCCCTATGGAAGCTTCTGAATTTGGCAGCGGACACCCACTCCC  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  63197  CTCCTGGAAATTCCATATCCCTATGGAAGCTTCTGAATTTGGCAGCGGACACCCACTCCC  63256

Query  241    TCTCCCGGCCCTTCTGATTGTGACCTGCTCTGGCTTAATGAAGCAGTTCAGCACTTCAGT  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  63257  TCTCCCGGCCCTTCTGATTGTGACCTGCTCTGGCTTAATGAAGCAGTTCAGCACTTCAGT  63316

Query  301    GGGGTAGAGGTTTGTGGAAAAAACAGTCCAGATGTTATCAAATACCTTTACATGTCTACT  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  63317  GGGGTAGAGGTTTGTGGAAAAAACAGTCCAGATGTTATCAAATACCTTTACATGTCTACT  63376

Query  361    TAGTAATAATTTGATTCAATACGTTAAGCATTCACTGTCTACTTCAATCTGGGTTAGATA  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  63377  TAGTAATAATTTGATTCAATACGTTAAGCATTCACTGTCTACTTCAATCTGGGTTAGATA  63436

Query  421    CATGAAATAACAGAAATACTCATCACA  447
              |||||||||||||||||||||||||||
Sbjct  63437  CATGAAATAACAGAAATACTCATCACA  63463
```

B)

Homo sapiens ecto-NOX disulfide-thiol exchanger 1 (ENOX1), RefSeqGene on chromosome 13

Sequence ID: NG_052843.1 **Length:** 580456 **Number of Matches:** 1

Range 1: 318016 to 318374 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
642 bits(711)	0.0	359/360(99%)	1/360(0%)	Plus/Minus

```
Query  1      TCTCCTAACTCCAGCATACGCACTCATCTCTGGGCTATCAATAGAAGCTAAAGAAAAAAT  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  318374 TCTCCTAACTCCAGCATACGCACTCATCTCTGGGCTATCAATAGAAGCTAAAGAAAAAAT  318315

Query  61     TAAAGAAACACTGGCTGCATGCCAAATTTTGGGTTAATAGGGAAGAGAAAGCTGGTTGGC  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  318314 TAAAGAAACACTGGCTGCATGCCAAATTTTGGGTTAATAGGGAAGAGAAAGCTGGTTGGC  318255

Query  121    AATCTCCCTTCCTAGTCCTTATCTTTTCCTCCCCTTATCTCTCCTGTGCTAGGGGTGCTG  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  318254 AATCTCCCTTCCTAGTCCTTATCTTTTCCTCCCCTTATCTCTCCTGTGCTAGGGGTGCTG  318195

Query  181    GGATGTCTGTACTCTGTAACCTCTTTCTCTGAATGGAGACAGAAGCAGGAAGTCAGGTGC  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  318194 GGATGTCTGTACTCTGTAACCTCTTTCTCTGAATGGAGACAGAAGCAGGAAGTCAGGTGC  318135

Query  241    CCTGCACTTTAATGTCTTATTATTCTTCTCCTTGAGTAAATCCTAACTGTTAACGGGAAA  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  318134 CCTGCACTTTAATGTCTTATTATTCTTCTCCTTGAGTAAATCCTAACTGTTAACGGGAAA  318075

Query  301    TAAAATGCTATATGTAAATGAACATCGTCACTTGGAAATCATCATTCTGGGGCTACAGCA  360
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  318074 TAAAATGCTATATGTAAATGAACATCGTCACTTGGAAATCATCATTCT-GGGCTACAGCA  318016
```

C)

Homo sapiens chromosome 8 clone RP11-420F14 map p21-p11, complete sequence

Sequence ID: AF276759.4 **Length:** 213611 **Number of Matches:** 1

Range 1: 167979 to 168385 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
730 bits(809)	0.0	406/407(99%)	0/407(0%)	Plus/Plus

```
Query  1       TGCCTTCCCTGCCCCTCCACTCCTCTGCCCAAAGACACACCACCAGCCTCCAGGGCCAGG  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  167979  TGCCTTCCCTGCCCCTCCACTCCTCTGCCCAAAGACACACCACCAGCCTCCAGGGCCAGG  168038

Query  61      CTCCAACAACCTCCTCTCTAGAAGGTTTATTGGACTGTCCTCTAACCTGGCCTGCGATTT  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  168039  CTCCAACAACCTCCTCTCTAGAAGGTTTATTGGACTGTCCTCTAACCTGGCCTGCGATTT  168098

Query  121     CTTTTCCAATATCCATCCCTCCTGCTTCACTGTTGCCTCCCTAGTCCTTCTCTTCTGGAG  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  168099  CTTTTCCAATATCCATCCCTCCTGCTTCACTGTTGCCTCCCTAGTCCTTCTCTTCTGGAG  168158

Query  181     AAAATGAACTCTACTCCTGCCTCCAGGATGGGACACACACACTGGCCTGAGCTATTCAGC  240
               |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Sbjct  168159  AAAATGAACTCTACTCCTGCCTCCAGGATGGGACACACACACAGGCCTGAGCTATTCAGC  168218

Query  241     ACAAGGCATTCTCTTTTCTCAGGGATTGGTTTAGAAATAGGCATGTGACCCACTTAGAGC  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  168219  ACAAGGCATTCTCTTTTCTCAGGGATTGGTTTAGAAATAGGCATGTGACCCACTTAGAGC  168278

Query  301     CAAAGAAACACCATGAGAGCTTTGCTGAGGTTTCTaaaaaaaGTTCACTCTTCGGAAAGC  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  168279  CAAAGAAACACCATGAGAGCTTTGCTGAGGTTTCTAAAAAAAGTTCACTCTTCGGAAAGC  168338

Query  361     AGGCCCGGAAATGGCTATCCCCCTGTTCCTCTAGACCAGTGATTCTC  407
               |||||||||||||||||||||||||||||||||||||||||||||
Sbjct  168339  AGGCCCGGAAATGGCTATCCCCCTGTTCCTCTAGACCAGTGATTCTC  168385
```

D)

Homo sapiens BAC clone RP11-557N21 from chromosome 2, complete sequence

Sequence ID: AC009242.6 **Length:** 211834 **Number of Matches:** 1

Range 1: 69744 to 70249 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
894 bits(991)	0.0	504/507(99%)	2/507(0%)	Plus/Minus

```
Query  1       GAATGGAACGGTCACTAATAACTTCCATATCATCAGTAGATGAGTGGTAGGAAAA-GGCC  59
               |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Sbjct  70249   GAATGGAACGGTCACTAATAACTTCCATATCATCAGTAGATGAGTGGTAGGAAAAAGGCC  70190

Query  60      TAGATAACTTTTTGTTCAACTTATAAAAGATATCACAAATGCAAGAAAAACATTAAACTT  119
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  70189   TAGATAACTTTTTGTTCAACTTATAAAAGATATCACAAATGCAAGAAAAACATTAAACTT  70130

Query  120     TTATTTTCCATTTTACCCACCTTTTTCCTCCACTCTCGCCTTCTTTTACCTTCCATGGCT  179
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  70129   TTATTTTCCATTTTACCCACCTTTTTCCTCCACTCTCGCCTTCTTTTACCTTCCATGGCT  70070

Query  180     TGAAGCTCATCAGGGAGGAGAAGAAAAGGACTAGCTGGCACAGAATTTCTTTCCCTGAAG  239
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  70069   TGAAGCTCATCAGGGAGGAGAAGAAAAGGACTAGCTGGCACAGAATTTCTTTCCCTGAAG  70010

Query  240     CTGTGCTTCATCTGTTAATCAAAATACTAACTTGAAAAGATTCTAAACAACCAGAATTCC  299
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  70009   CTGTGCTTCATCTGTTAATCAAAATACTAACTTGAAAAGATTCTAAACAACCAGAATTCC  69950

Query  300     AGTAAGAAGGAGAAACTTCAGTACGAAAATCAGGATGTCTTATGACGCAAGTGGTCCTAT  359
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  69949   AGTAAGAAGGAGAAACTTCAGTACGAAAATCAGGATGTCTTATGACGCAAGTGGTCCTAT  69890

Query  360     ACAGACCAATCATTTGTTTCTGGAACAAGAGGTGCTGTCAGGAAAACAGTTGGCATTAAG  419
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  69889   ACAGACCAATCATTTGTTTCTGGAACAAGAGGTGCTGTCAGGAAAACAGTTGGCATTAAG  69830

Query  420     TCAAGGAAAGGTGAAAGAAAGCCCAACTCCATCACCTCCAAGCACTCTGAAGGTAACTCT  479
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  69829   TCAAGGAAAGGTGAAAGAAAGCCCAACTCCATCACCTCCAAGCACTCTGAAGGTAACTCT  69770

Query  480     TGCTCCTGTAAAGGAggggggggACCA  506
               |||||||||||||| || ||||||||
Sbjct  69769   TGCTCCTGTAAAGGA-GGAGGGGACCA  69744
```

E)

Homo sapiens FOSMID clone ABC16-1602J11 from chromosome 16, complete sequence

Sequence ID: AC238650.2 **Length:** 37234 **Number of Matches:** 1

Range 1: 9614 to 10055 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
817 bits(442)	0.0	442/442(100%)	0/442(0%)	Plus/Minus

```
Query  1      GGGGGCCTGGGGAGCCGCCTGAGCCCCCGATGCTGGGATGCGGCGGTGGGGCCGGCTCAC  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  10055  GGGGGCCTGGGGAGCCGCCTGAGCCCCCGATGCTGGGATGCGGCGGTGGGGCCGGCTCAC  9996

Query  61     CCAGGCAGAGAGGAGGGTGTAGACCCCTGGCCAGCGCCCGCAGCTGTAGCTCTTGACCCA  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  9995   CCAGGCAGAGAGGAGGGTGTAGACCCCTGGCCAGCGCCCGCAGCTGTAGCTCTTGACCCA  9936

Query  121    GCTCATGACGCCAACCTGCCCTCAGGAACCATGCGCCTGGCAGACCAGAGGCCCTCCCGA  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  9935   GCTCATGACGCCAACCTGCCCTCAGGAACCATGCGCCTGGCAGACCAGAGGCCCTCCCGA  9876

Query  181    GTCACCCTGAGGGCAGAGAAGGCAAGGACTAGGAGGTGGGGGGTCCCGTTTCTGGGCTCC  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  9875   GTCACCCTGAGGGCAGAGAAGGCAAGGACTAGGAGGTGGGGGGTCCCGTTTCTGGGCTCC  9816

Query  241    GGCACCCCCACGCTCCGTGGCCAGCCTGGCTCATCAGGTTGCTCTCGTCGGGGAACCGCA  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  9815   GGCACCCCCACGCTCCGTGGCCAGCCTGGCTCATCAGGTTGCTCTCGTCGGGGAACCGCA  9756

Query  301    GAGTGAAGCCGCAGGGAGAACCCACTACACACCCACCGAAACGGCCAGAAGCTTTTTAAA  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  9755   GAGTGAAGCCGCAGGGAGAACCCACTACACACCCACCGAAACGGCCAGAAGCTTTTTAAA  9696

Query  361    CAGTCAAGCTCTGCTCACATCTCCACCCCACCCCTTCCCTTTCCCTCCAATAGATTCTGT  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  9695   CAGTCAAGCTCTGCTCACATCTCCACCCCACCCCTTCCCTTTCCCTCCAATAGATTCTGT  9636

Query  421    CCTAAAGCCTAGCATGGGGCAG  442
              ||||||||||||||||||||||
Sbjct  9635   CCTAAAGCCTAGCATGGGGCAG  9614
```

F)

Homo sapiens BAC clone RP11-622P13 from 7, complete sequence

Sequence ID: AC073846.6 **Length:** 123004 **Number of Matches:** 2

Range 1: 92119 to 92581 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
856 bits(463)	0.0	463/463(100%)	0/463(0%)	Plus/Plus

```
Query  1      GGTGTCCATCCTCACCTGCCTCCGATCATTTTCAGGCTGGAGCTGAACACAGAGGGGTCT  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  92119  GGTGTCCATCCTCACCTGCCTCCGATCATTTTCAGGCTGGAGCTGAACACAGAGGGGTCT  92178

Query  61     CACTCCCTGCCCCTCTAAAGAAGAATCAGCTCTGTGCAAGACAAGGTGTGTGGGAGGCCT  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  92179  CACTCCCTGCCCCTCTAAAGAAGAATCAGCTCTGTGCAAGACAAGGTGTGTGGGAGGCCT  92238

Query  121    TCTGAAGGACCCCATCTGGTTCCCATGGTAACACTTCTTCCTAGTCCCTCTCTTGTCCAT  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  92239  TCTGAAGGACCCCATCTGGTTCCCATGGTAACACTTCTTCCTAGTCCCTCTCTTGTCCAT  92298

Query  181    GTCTCAACTTTCTCTGCTTCTTACCTACCCTGAGGAGGGGAGGCTCAACCTGGTAGCTGC  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  92299  GTCTCAACTTTCTCTGCTTCTTACCTACCCTGAGGAGGGGAGGCTCAACCTGGTAGCTGC  92358

Query  241    TGCCTTCTAAAGGCTGCTCCCCAATTCATGATAGTGGAGAAAGGTCCAGACCTGCCCAAC  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  92359  TGCCTTCTAAAGGCTGCTCCCCAATTCATGATAGTGGAGAAAGGTCCAGACCTGCCCAAC  92418

Query  301    CTTTGCCCCTTTTTACCTTTTGGTTGCAAACTCTACACCATGTCAATCATGAAGGAGAGC  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  92419  CTTTGCCCCTTTTTACCTTTTGGTTGCAAACTCTACACCATGTCAATCATGAAGGAGAGC  92478

Query  361    CaaaaaaaaaaaaaaaaaGTGGGGACCGGGCACTGTGGCTTACACCTGTGATCCCAGCA  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  92479  CAAAAAAAAAAAAAAAAAGTGGGGACCGGGCACTGTGGCTTACACCTGTGATCCCAGCA  92538

Query  421    CTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGATGG  463
              |||||||||||||||||||||||||||||||||||||||||||
Sbjct  92539  CTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGATGG  92581
```

G)

Homo sapiens dedicator of cytokinesis 3 (DOCK3), RefSeqGene on chromosome 3

Sequence ID: NG_028012.2 **Length:** 715958 **Number of Matches:** 1

Range 1: 672436 to 672721 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
529 bits(286)	9e-148	286/286(100%)	0/286(0%)	Plus/Plus

```
Query  1       TGCATGAGCATTGATGGTCCCCTATTCATCCCACATTATAGTAATGCATAGAAAAGCAGG  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  672436  TGCATGAGCATTGATGGTCCCCTATTCATCCCACATTATAGTAATGCATAGAAAAGCAGG  672495

Query  61      ATTAAATTTTAACATGGACTAGGGAGTAAAACACAAAGTGTGCCTGAGTACTGTAGATGG  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  672496  ATTAAATTTTAACATGGACTAGGGAGTAAAACACAAAGTGTGCCTGAGTACTGTAGATGG  672555

Query  121     ACAAGAAGGAGAGAAGGACTAGGAGGCTCCATAGGCCACCTTCTCAGGAATCTCCTCACT  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  672556  ACAAGAAGGAGAGAAGGACTAGGAGGCTCCATAGGCCACCTTCTCAGGAATCTCCTCACT  672615

Query  181     TCTGCAATCCTGCCTTTGCCCTGAATTTCTGGAGGAGCAGAGTCTATGCTGAGTAGAGAC  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  672616  TCTGCAATCCTGCCTTTGCCCTGAATTTCTGGAGGAGCAGAGTCTATGCTGAGTAGAGAC  672675

Query  241     CACTCAGTGAGCCCCCTCATGAGCTCATCCTCAGAGGGGCAGTTAT  286
               ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  672676  CACTCAGTGAGCCCCCTCATGAGCTCATCCTCAGAGGGGCAGTTAT  672721
```

H)

Homo sapiens sarcoglycan zeta (SGCZ), RefSeqGene (LRG_208) on chromosome 8

Sequence ID: NG_008899.1 **Length:** 1155420 **Number of Matches:** 1

Range 1: 401680 to 402179 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
924 bits(500)	0.0	500/500(100%)	0/500(0%)	Plus/Plus

```
Query  1       CGCCGAAATAAATTACTCATCTGGTATAAAATGTATTTTTTCTTATGGAAACCTACTATT  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  401680  CGCCGAAATAAATTACTCATCTGGTATAAAATGTATTTTTTCTTATGGAAACCTACTATT  401739

Query  61      TAAAATTATCCCAAGCTATTAAATTTCTTGCTCCAGAGTCTGAAACTCCTTACTAACCAA  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  401740  TAAAATTATCCCAAGCTATTAAATTTCTTGCTCCAGAGTCTGAAACTCCTTACTAACCAA  401799

Query  121     TCATGCAATTATGTGGAGAGTAAGATTCTCAGAAAGGATAGAATTTTAGTTCTTTAAATA  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  401800  TCATGCAATTATGTGGAGAGTAAGATTCTCAGAAAGGATAGAATTTTAGTTCTTTAAATA  401859

Query  181     CCATGACTATGAATCCTCCCATTACAACATTTCCCTGAATGCAAGGTGAAATCCAATAAA  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  401860  CCATGACTATGAATCCTCCCATTACAACATTTCCCTGAATGCAAGGTGAAATCCAATAAA  401919

Query  241     TGAAATTCTAGATAGTAGCCTTTTGAAACTGCCTATTTATCTCTTATGTTAGGTTAGCCT  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  401920  TGAAATTCTAGATAGTAGCCTTTTGAAACTGCCTATTTATCTCTTATGTTAGGTTAGCCT  401979

Query  301     GAGAGCTTCTTTGGAGCCAGTGCTTGTAGATGCAGCAGAAGAGAAGGACCAGGTGGATTG  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  401980  GAGAGCTTCTTTGGAGCCAGTGCTTGTAGATGCAGCAGAAGAGAAGGACCAGGTGGATTG  402039

Query  361     CTCTTTAGTTTACATTTGATGGTTGATACAGCCCATAACATTTAGTCTGTTTTTAACATC  420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  402040  CTCTTTAGTTTACATTTGATGGTTGATACAGCCCATAACATTTAGTCTGTTTTTAACATC  402099

Query  421     TTTTGGGTTTAATGTTTCAGACATGCAGCATTTAAACTTAAATAACCTAGTTTTGTAACT  480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  402100  TTTTGGGTTTAATGTTTCAGACATGCAGCATTTAAACTTAAATAACCTAGTTTTGTAACT  402159

Query  481     TGAttttttttttGTTCTTC  500
               |||||||||||||||||||
Sbjct  402160  TGATTTTTTTTTTGTTCTTC  402179
```

I)

Homo sapiens BAC clone RP11-693G13 from 2, complete sequence

Sequence ID: AC064837.6 **Length:** 111857 **Number of Matches:** 1

Range 1: 92249 to 92690 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
817 bits(442)	0.0	442/442(100%)	0/442(0%)	Plus/Plus

```
Query  1      GATTATGATTCTCCACTTCTTCTCTTCCTGCTCTCTCAGTAAATCATTGCACACAGGTGT  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  92249  GATTATGATTCTCCACTTCTTCTCTTCCTGCTCTCTCAGTAAATCATTGCACACAGGTGT  92308

Query  61     ACATGAGCAATTTATAATATACTGATCATTTTAAAGAATAAACTGAAGAGAGGGTTTCAT  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  92309  ACATGAGCAATTTATAATATACTGATCATTTTAAAGAATAAACTGAAGAGAGGGTTTCAT  92368

Query  121    ATCAAATACTCCAAGGCAAACTCTTCTAAAATCATCTTGAGACATATCTCTTCTTTCAGG  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  92369  ATCAAATACTCCAAGGCAAACTCTTCTAAAATCATCTTGAGACATATCTCTTCTTTCAGG  92428

Query  181    TCTCTTTTACCTCTCACGTGATCAAACAGATGttttttttcttcctcttttcttccccccag  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  92429  TCTCTTTTACCTCTCACGTGATCAAACAGATGTTTTTTTCTTCCTCTTTTCTTCCCCCAG  92488

Query  241    acactgctcagcttctgcttagtccttctcttcttctcttccAAAAGAAAAAGTGTTTCT  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  92489  ACACTGCTCAGCTTCTGCTTAGTCCTTCTCTTCTTCTCTTCCAAAAGAAAAAGTGTTTCT  92548

Query  301    TTTCTTTTACACGTGTAATAATCCTACCATGTCTCCATATCCTAGATGACAATTGATAAC  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  92549  TTTCTTTTACACGTGTAATAATCCTACCATGTCTCCATATCCTAGATGACAATTGATAAC  92608

Query  361    AGACACCCCCAACCCCTGGCCTCTACCATGAACAAGTTTATTGAAACAGCAAATGTAGGC  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  92609  AGACACCCCCAACCCCTGGCCTCTACCATGAACAAGTTTATTGAAACAGCAAATGTAGGC  92668

Query  421    CTTTCCTTCTAGAAGATTTGTA  442
              |||||||||||||||||||||||
Sbjct  92669  CTTTCCTTCTAGAAGATTTGTA  92690
```

J)

Homo sapiens 3 BAC RP11-31B16 (Roswell Park Cancer Institute Human BAC Library) complete sequence

Sequence ID: AC117509.9 **Length:** 113201 **Number of Matches:** 1

Range 1: 87513 to 87989 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
874 bits(473)	0.0	476/477(99%)	1/477(0%)	Plus/Minus

```
Query  1      TCATCCACAATGGGAAATGAGCAAGTGGTTCCAAAAGCAGAAGTGGGTATGGGTACACAT  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  87989  TCATCCACAATGGGAAATGAGCAAGTGGTTCCAAAAGCAGAAGTGGGTATGGGTACACAT  87930

Query  61     GGGAGTGTATAAAGCTCAAAGTCTCCTTATACACGGCATTGATTCAGAAAGGAACTGACA  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  87929  GGGAGTGTATAAAGCTCAAAGTCTCCTTATACACGGCATTGATTCAGAAAGGAACTGACA  87870

Query  121    CTGTTGATGAAATTTATCACACAAGATGGGACGTGAACAAATCCTTCATCCTAAAAAGGC  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  87869  CTGTTGATGAAATTTATCACACAAGATGGGACGTGAACAAATCCTTCATCCTAAAAAGGC  87810

Query  181    AGAAACTCAGAAGAGAATGACTAGGGAGAATGAGTCACTAGCAGCTGTTAAGAATCATAG  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  87809  AGAAACTCAGAAGAGAATGACTAGGGAGAATGAGTCACTAGCAGCTGTTAAGAATCATAG  87750

Query  241    TTAGATCTCTATTCCTGATTTAACTCTCTGCTTTATCTACCCCAGAGGCCAGTCGTTCTC  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  87749  TTAGATCTCTATTCCTGATTTAACTCTCTGCTTTATCTACCCCAGAGGCCAGTCGTTCTC  87690

Query  301    AAACTTCAGCTGCATTCAGATCACTTCAAGGGCTTGTTAAAACAGGATTGCTGATCTACA  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  87689  AAACTTCAGCTGCATTCAGATCACTTCAAGGGCTTGTTAAAACAGGATTGCTGATCTACA  87630

Query  361    CACCCAGAATTTCTGATTCAGTACATGTGGGTGGGGCAAGAGAATTTTTAGCTCTAACAA  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  87629  CACCCAGAATTTCTGATTCAGTACATGTGGGTGGGGCAAGAGAATTTTTAGCTCTAACAA  87570

Query  421    GTTCCCAGGTGATGCTGGTCGGGGATACCACACCATAAGACCCACT-CTTTAGGCAA  476
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Sbjct  87569  GTTCCCAGGTGATGCTGGTCGGGGATACCACACCATAAGACCCACTGCTTTAGGCAA  87513
```

K)

Homo sapiens T cell receptor alpha delta locus (TCRA/TCRD) on chromosome 14

Sequence ID: NG_001332.3 **Length:** 930229 **Number of Matches:** 1

Range 1: 868017 to 868472 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
837 bits(453)	0.0	455/456(99%)	0/456(0%)	Plus/Plus

```
Query  1       CCTTGAACATTCCTGAAGTgggggggCAGGCGATGTCACTTGCAGTGACAGCTGGCCTCT  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  868017  CCTTGAACATTCCTGAAGTGGGGGGGCAGGCGATGTCACTTGCAGTGACAGCTGGCCTCT  868076

Query  61      CTAGCTCACCCCCACATCTGTGCCAACCAGAGAAGCTGGGCCCTTTGTAGCACAAGCCTC  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  868077  CTAGCTCACCCCCACATCTGTGCCAACCAGAGAAGCTGGGCCCTTTGTAGCACAAGCCTC  868136

Query  121     AGGAGCAGAGAATGGAGGGACATTCTCAACTGAGAGGAGAGGGACTAGGAGGCTAGATGG  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  868137  AGGAGCAGAGAATGGAGGGACATTCTCAACTGAGAGGAGAGGGACTAGGAGGCTAGATGG  868196

Query  181     GAAATGAGGTGACTCTACAGAAAATCGAATATGGGATGGTATAAAAGGATGAGAAAGATG  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  868197  GAAATGAGGTGACTCTACAGAAAATCGAATATGGGATGGTATAAAAGGATGAGAAAGATG  868256

Query  241     GCAGGGGGTCCACGATCGCTCTTGTATTTAGACCTGGGGGATGCACTGATAGATACCACC  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  868257  GCAGGGGGTCCACGATCGCTCTTGTATTTAGACCTGGGGGATGCACTGATAGATACCACC  868316

Query  301     TGGAGAGAGAAAGAGGATTTCTGTTTCCCTCTGAAGAAAAATGTGGCAGAAAAATCAAAA  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  868317  TGGAGAGAGAAAGAGGATTTCTGTTTCCCTCTGAAGAAAAATGTGGCAGAAAAATCAAAA  868376

Query  361     CAGATCGCACATGTCAAGCCTCTCCTAATTACAGCCTTTCATTCTGGTTTCTCTCAGCTG  420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  868377  CAGATCGCACATGTCAAGCCTCTCCTAATTACAGCCTTTCATTCTGGTTTCTCTCAGCTG  868436

Query  421     TTCCAGCAGGAAATGGTAGGAAAGGGCCATTGGTGC  456
               |||||||||||||||||||||||| |||||||||
Sbjct  868437  TTCCAGCAGGAAATGGTAGGAAAGGGTCATTGGTGC  868472
```

L)

Homo sapiens chromosome 8, clone RP11-238G17, complete sequence

Sequence ID: AC021955.11 **Length:** 192139 **Number of Matches:** 1

Range 1: 132085 to 132582 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
920 bits(498)	0.0	498/498(100%)	0/498(0%)	Plus/Minus

```
Query  1       ACTTGTCCATCCCCTCTCTCACTCTATCTTGGCCCCATGTCTTGCTTTTCACCACTCTCG  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132582  ACTTGTCCATCCCCTCTCTCACTCTATCTTGGCCCCATGTCTTGCTTTTCACCACTCTCG  132523

Query  61      CAGGGCTTCCCTTCTCCAGAACATGATTTGTAGAGTTGTTTATGCTTTGTTTGCTGAGTC  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132522  CAGGGCTTCCCTTCTCCAGAACATGATTTGTAGAGTTGTTTATGCTTTGTTTGCTGAGTC  132463

Query  121     TCCTGCATACTCTGATTTTACTCTCACCTTTGCAGACCATTCCAGTCATAGCCATCTCTC  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132462  TCCTGCATACTCTGATTTTACTCTCACCTTTGCAGACCATTCCAGTCATAGCCATCTCTC  132403

Query  181     CCTTGTCCTTCTCATCCCCGACTACAGTCTCTCCAACACAATGAAACACTGAGGTTGAAA  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132402  CCTTGTCCTTCTCATCCCCGACTACAGTCTCTCCAACACAATGAAACACTGAGGTTGAAA  132343

Query  241     GAAGGATTCTAGGCACAAGTGAAGTTTGGACATAACTAAGAATGGGGTGGTGGAAGAAGT  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132342  GAAGGATTCTAGGCACAAGTGAAGTTTGGACATAACTAAGAATGGGGTGGTGGAAGAAGT  132283

Query  301     CATTGCCAAACAGATCCAGAAGGGGATGTCTGGGAAGTCCATGAAACAAAGTCTTCTCCA  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132282  CATTGCCAAACAGATCCAGAAGGGGATGTCTGGGAAGTCCATGAAACAAAGTCTTCTCCA  132223

Query  361     AACTCTAAGCACCATCTCTATCTCCCTGGTGTCTTTATGCACAGAGATTCCAGTTCTTTA  420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132222  AACTCTAAGCACCATCTCTATCTCCCTGGTGTCTTTATGCACAGAGATTCCAGTTCTTTA  132163

Query  421     CTTTGCCTTGGCTTTCATCAAAATGTCCTATGTTGATTCTCCCTTTCCTCTCTTCTCACC  480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  132162  CTTTGCCTTGGCTTTCATCAAAATGTCCTATGTTGATTCTCCCTTTCCTCTCTTCTCACC  132103

Query  481     TTCCTTCATACAAACCCC  498
               ||||||||||||||||||
Sbjct  132102  TTCCTTCATACAAACCCC  132085
```

M)

Human DNA sequence from clone RP11-416N2 on chromosome 10, complete sequence

Sequence ID: AL121929.17 **Length:** 203262 **Number of Matches:** 1

Range 1: 112667 to 113119 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
837 bits(453)	0.0	453/453(100%)	0/453(0%)	Plus/Plus

```
Query  1       CAGAAGGTCTCTACTGAAAACATGTGGGTCTGATTCTGTGCCTGGCACGGGTGGGGGTCG  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  112667  CAGAAGGTCTCTACTGAAAACATGTGGGTCTGATTCTGTGCCTGGCACGGGTGGGGGTCG  112726

Query  61      GAGACAGGTTGAGGTGGAGATTATAAGATGGCATGAGGCCTCTCCAGGAGCTGCAGAGCA  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  112727  GAGACAGGTTGAGGTGGAGATTATAAGATGGCATGAGGCCTCTCCAGGAGCTGCAGAGCA  112786

Query  121     GTGGATGGATGAGAAGGGCGGCTGCTTGCAGGGGGACTTGGCACGAGACTGCTTGTGCTG  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  112787  GTGGATGGATGAGAAGGGCGGCTGCTTGCAGGGGGACTTGGCACGAGACTGCTTGTGCTG  112846

Query  181     GGTGCGGCAGGGCGACCCAGACAGCCCCATGCAGTGGACATGGCGGTGGGTGATACAATG  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  112847  GGTGCGGCAGGGCGACCCAGACAGCCCCATGCAGTGGACATGGCGGTGGGTGATACAATG  112906

Query  241     CTTGCCAAGGAACCAGAGATATCCAGAGATGAGAATGACTAGGCAGAAAAGAATTCTTAG  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  112907  CTTGCCAAGGAACCAGAGATATCCAGAGATGAGAATGACTAGGCAGAAAAGAATTCTTAG  112966

Query  301     AGGAGAAGAGACTTGAATTGAACCTCAGAGTGACAGGCTGCATGGGCTCACTCTAATGGC  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  112967  AGGAGAAGAGACTTGAATTGAACCTCAGAGTGACAGGCTGCATGGGCTCACTCTAATGGC  113026

Query  361     TTCATAGGGGATTATTCCAGATGAAAATGAGGCATTCAAGGCCGGGCATGGTGGCTCATG  420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  113027  TTCATAGGGGATTATTCCAGATGAAAATGAGGCATTCAAGGCCGGGCATGGTGGCTCATG  113086

Query  421     CCTGTAATCCTGGCACTTTGGGAGGCTGAGGCA  453
               |||||||||||||||||||||||||||||||
Sbjct  113087  CCTGTAATCCTGGCACTTTGGGAGGCTGAGGCA  113119
```

N)

Human DNA sequence from clone RP1-221C16 on chromosome 6, complete sequence

Sequence ID: AL353759.8 **Length:** 101099 **Number of Matches:** 1

Range 1: 89003 to 89472 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
869 bits(470)	0.0	470/470(100%)	0/470(0%)	Plus/Plus

```
Query  1      TGATTAAATAAATACAAATAAAGCACATTAGCTATTATTCTGAACACTGCTACATTAGCT  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89003  TGATTAAATAAATACAAATAAAGCACATTAGCTATTATTCTGAACACTGCTACATTAGCT  89062

Query  61     ACATTGAAAATTAAAAAAGATGTTAAACAGATTTTAAGGTTTTCTCGTCATCACTGAAGT  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89063  ACATTGAAAATTAAAAAAGATGTTAAACAGATTTTAAGGTTTTCTCGTCATCACTGAAGT  89122

Query  121    GAGAAAAACAAATTATTTTGTGAAACAAAGATAAGACACAATAAATTCAGAGTACACTTT  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89123  GAGAAAAACAAATTATTTTGTGAAACAAAGATAAGACACAATAAATTCAGAGTACACTTT  89182

Query  181    TGTTAGATGCTAACCCACAGAGCGGAATACTTTTTGCGCAGCACCTTCATTCCTGGAGAC  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89183  TGTTAGATGCTAACCCACAGAGCGGAATACTTTTTGCGCAGCACCTTCATTCCTGGAGAC  89242

Query  241    TTAGGGCCCGGGAAATGGGAATGGTGGCAGGAGAGGAGAAGGGCTAGGGGGCTGGGGAGC  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89243  TTAGGGCCCGGGAAATGGGAATGGTGGCAGGAGAGGAGAAGGGCTAGGGGGCTGGGGAGC  89302

Query  301    GTGGGGTGAAGGAGTGGGGGCTTAAAAAGGACAGTCAAATTTAGGAAGAACAGTTTCTTA  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89303  GTGGGGTGAAGGAGTGGGGGCTTAAAAAGGACAGTCAAATTTAGGAAGAACAGTTTCTTA  89362

Query  361    AACTGAAGTTAGCTAAGCAGCGGGCTTCAGCCAAAGATAATAATACTAACTTGAAAGTAA  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89363  AACTGAAGTTAGCTAAGCAGCGGGCTTCAGCCAAAGATAATAATACTAACTTGAAAGTAA  89422

Query  421    AGCTGTTTGTTAAATAGGCTTTCATTTTAATTTTTTAAAAAATATTTTCA  470
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  89423  AGCTGTTTGTTAAATAGGCTTTCATTTTAATTTTTTAAAAAATATTTTCA  89472
```

O)

Homo sapiens chromosome 14 clone ABC9_3_5_000046182700_E23, complete sequence

Sequence ID: AC279153.1 **Length:** 41380 **Number of Matches:** 1

Range 1: 11351 to 11821 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
843 bits(456)	0.0	467/472(99%)	1/472(0%)	Plus/Plus

```
Query  1      TTCAATTTTACAAATATGGTCATAGACATTATCATTATCCACATGGATAACAAGTCAATT  60
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Sbjct  11351  TTCAATTTTACAAATATGGTCATAGACATTATCATTATCCACATAGATAACAAGTCAATT  11410

Query  61     ACCCTCAAAATATCCTCTTGTTCTGTAATTCCTCCTTCCTAGACCTTCCCTTCTCCTACA  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  11411  ACCCTCAAAATATCCTCTTGTTCTGTAATTCCTCCTTCCTAGACCTTCCCTTCTCCTACA  11470

Query  121    ATATTGACAGTGAACTACTGATTTTTATGTAACTTTAGATTACTTATCAATACATCAGGT  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  11471  ATATTGACAGTGAACTACTGATTTTTATGTAACTTTAGATTACTTATCAATACATCAGGT  11530

Query  181    AATAAAAGTTATAGATTTATGTGTGTTGTGGGGTGGCTGTATATAAGTTTCTGTGTgaga  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  11531  AATAAAAGTTATAGATTTATGTGTGTTGTGGGGTGGCTGTATATAAGTTTCTGTGTGAGA  11590

Query  241    gagagaaggagggaggaaggaaggcagaaaaagagaggaaTCCTACATAATTGACCACAA  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  11591  GAGAGAAGGAGGGAGGAAGGAAGGCAGAAAAAGAGAGGAATCCTACATAATTGACCACAA  11650

Query  301    TTTATGAGGTTCTCAAGTAATTATGGGGAATTAGTCCTTACAGACAAGGCTGATATAAGA  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  11651  TTTATGAGGTTCTCAAGTAATTATGGGGAATTAGTCCTTACAGACAAGGCTGATATAAGA  11710

Query  361    TGGAGAGGACAACTTGACACACCTAGCTATGGTTATATATTTATATCAATATCATTTTCT  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  11711  TGGAGAGGACAACTTGACACACCTAGCTATGGTTATATATTTATATCAATATCATTTTCT  11770

Query  421    AATCATACAAACACATGCATTAAAATAAAGGTAGTGGAGGGGGTCTGGGTGG  472
              |||||||||||||||||||| |||| |||||||||||||| |||||| |||
Sbjct  11771  AATCATACAAACACATGCATTAGAATAGAGGTAGTGGAGGGTGTCTGG-TGG  11821
```

P)

Homo sapiens chromosome 1 clone RP11-330M19, complete sequence

Sequence ID: AC096541.2 **Length:** 200368 **Number of Matches:** 1

Range 1: 166029 to 166545 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
942 bits(510)	0.0	515/517(99%)	2/517(0%)	Plus/Plus

```
Query  1       AGGCCCAGCCAGGGAGGTGCCCAGTTCTGGAGTGGAGGCTGTGGGGAACAGCACACCCTT  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  166029  AGGCCCAGCCAGGGAGGTGCCCAGTTCTGGAGTGGAGGCTGTGGGGAACAGCACACCCTT  166088

Query  61      CCCCTTCTCCCTGCCAAAAGGCAGATCCTGAGCCATGGTGCGCCAAGGTTTAGGGTCCAA  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  166089  CCCCTTCTCCCTGCCAAAAGGCAGATCCTGAGCCATGGTGCGCCAAGGTTTAGGGTCCAA  166148

Query  121     TTTCCCAGATGCCTTTCAGTGTCTGATAGGAGATGCCAGAGAAGAGAAAGACCAGGAGGC  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  166149  TTTCCCAGATGCCTTTCAGTGTCTGATAGGAGATGCCAGAGAAGAGAAAGACCAGGAGGC  166208

Query  181     CAGGTGATTTATGAGGGAAATCTTGGCTTTGACAAAGCTTTAGCCCTAGAGATTAAGGCA  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  166209  CAGGTGATTTATGAGGGAAATCTTGGCTTTGACAAAGCTTTAGCCCTAGAGATTAAGGCA  166268

Query  241     GGCGAAAGCAGAGTGCACAGCCTCTTCCTGAGGGGTGTTGGGGTTTCTCTGGGCATCTGG  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  166269  GGCGAAAGCAGAGTGCACAGCCTCTTCCTGAGGGGTGTTGGGGTTTCTCTGGGCATCTGG  166328

Query  301     ATCTGGATCCCTGAGGATAGCGCTGAGCTCTTGGCTGACTCCAGCTACTCCTCAGCGCCA  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  166329  ATCTGGATCCCTGAGGATAGCGCTGAGCTCTTGGCTGACTCCAGCTACTCCTCAGCGCCA  166388

Query  361     CATCCTGGAGCAGAACAAGGCCCCTGGTGGCTTAGGAAGTGGGGAGAGAGGAGCTACGAG  420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  166389  CATCCTGGAGCAGAACAAGGCCCCTGGTGGCTTAGGAAGTGGGGAGAGAGGAGCTACGAG  166448

Query  421     TCCGCGATAGTAAACCAGGCCTGCTCTGCCTGCCCTCCTCCCTGCTACTGCCCTGCAGAG  480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  166449  TCCGCGATAGTAAACCAGGCCTGCTCTGCCTGCCCTCCTCCCTGCTACTGCCCTGCAGAG  166508

Query  481     GCTTCAAAGGACAGCAGGAGCCTCAGAA-TGGC-GTG  515
               |||||||||||||||||||||||||||| |||| |||
Sbjct  166509  GCTTCAAAGGACAGCAGGAGCCTCAGAACTGGCTGTG  166545
```

Figure S5: Nucleotide blast of sequencing data from the 16 off-target locations. The green labeled sequence marks the off-target location.

8.2 List of figures

8.3 List of tables