Stem Cell Reports, Volume 15

# **Supplemental Information**

# Precise Correction of Heterozygous SHOX2 Mutations in hiPSCs De-

rived from Patients with Atrial Fibrillation via Genome Editing and Sib

## Selection

Simon Alexander Sumer, Sandra Hoffmann, Svenja Laue, Birgit Campbell, Kristin Raedecke, Viktoria Frajs, Sebastian Clauss, Stefan Kääb, Johannes W.G. Janssen, Anna Jauch, Karl-Ludwig Laugwitz, Tatjana Dorn, Alessandra Moretti, and Gudrun A. Rappold



SHOX2 c.849C>A\_1 (46,XY)

SHOX2 c.\*28T>C\_1 (46,XY)

SHOX2 c.\*28T>C 3 (46,XY)

Figure S1 Generation and characterization of patient-specific hiPSCs from patient I (SHOX2 c.849C>A) and patient II (SHOX2 c.\*28T>C). Related to Figure 1 hiPSC were generated by reprogramming peripheral blood mononuclear cells (PBMCs) using the CytoTune-hiPSC 2.0 Sendai Reprogramming Kit (Life Technologies). (A) Integration-free reprogramming scheme with Sendai viruses, leading to the generation of two hiPSC clones from patient I and three hiPSC clones from patient II. (B) Sequencing of SHOX2 c.849C>A (= SHOX2 p.H283Q) and SHOX2 c.\*28T>C mutations in patient-derived hiPSC clones. (C) Detection of alkaline phosphatase activity as a pluripotency marker, representative bright field image of an hiPSC colony stained for alkaline phosphatase. (D) Confirmation of loss of Sendai viral transgenes after reprogramming by RT-PCR, a = uninfected PBMCs (negative control), b = infected PBMCs (positive control), c-e = hiPSC clones from patient I and patient II. (E) Immunofluorescence detection of the endogenous pluripotency markers NANOG and TRA-1-81 in patient-specific hiPSCs at passage 17 or 18 after reprogramming; scale bar, 50 μm. (F) Analysis of expression profile of endogenous pluripotency genes in patient-specific hiPSCs by qRT-PCR. Values are normalized to GAPDH and relative to uninfected parental patient PBMCs. (G) Schematic of the spontaneous differentiation of hiPSCs into EBs. qRT-PCR analysis of lineage markers specific for each of the three embryonic germlayers after 21 days of spontaneous EB differentiation in patient-specific hiPSCs. Values are normalized to GAPDH and relative to hiPSCs harvested at day 0. (H) M-FISH analysis of the two SHOX2 c.849C>A clones and three SHOX2 c.\*28T>C clones showing no chromosomal abnormalities.



**Figure S2 Allele quantification of SHOX2 c.849C>A gRNA-1 transfected sib-selections via NGS. Related to Figure 3** Alleles with SHOX2 c.849-spanning deletions cause an uncertainty in allele quantification that is addressed by defining those alleles as all WT or all Mut. The resulting span of possible allele ratios is represented as error bars. However, due to the high numbers of reads containing large deletions, the ratios cannot be determined precisely enough.

Abbreviations: Indel = Insertions/Deletions, N/A = non-assignable



**Figure S3 Single cell-cloning of hiPSCs. Related to Figure 4** (A) Coating free method for single-cell cloning via limited dilution. A single-cell suspension is generated from the chosen sib-selection and diluted to 10 cells/ml. The extracellular matrix is added to the cell suspension and 1 cell/well is seeded on 96-well plates (100 µl). (B) Single-cell derived hiPSC colony after 72h (upper row; scale bar, 50 µm) and 8 days (lower row; scale bar, 200 µm).



Figure S4 Re-characterization of SHOX2 c.849C>A\_isoWT and SHOX2 c.\*28T>C\_isoWT. Related to Figure 4 lsogenic control lines were generated as described above and re-characterized to confirm the preservation of pluripotency capacity and karyotype.

(A) Loss of heterozygous SHOX2 c.849C>A (= SHOX2 p.H283Q) and SHOX2 c.\*28T>C mutations in isogenic control clones. (B) Detection of alkaline phosphatase activity in isogenic hiPSCs as a pluripotency marker; scale bar (overview picture), 4 mm; scale bar (colony picture), 300 μm. (C) Immunofluorescence detection of the endogenous pluripotency markers OCT4 and SOX2 in isogenic hiPSCs; scale bar, 100 μm. (D) qRT-PCR analysis of pluripotency genes at d0 and germ layer markers after 21 days of EB differentiation (see the schematic above). (E) Giemsa banding of isogenic control lines revealing no chromosomal aberrations.

## **Supplemental Data Items**

### Patient I (SHOX c.849C>A)

The male patient with the *SHOX2* c.849C>A mutation was born in 1949 and developed AF at the age of 56 years. When he was recruited for study enrollment in 2009 at the age of 62, his AF had become persistent. An echocardiography at that time revealed a slightly enlarged left atrium (LA diameter = 43 mm) and a reduced EF of 43 %. In addition, the patient was diagnosed with arterial hypertension. The patient's ECG showed sinus rhythm with normal heartbeat rate and unremarkable P-, PR-, QRS- and QT-durations. No other family member developed AF (parents, one brother, one sister, one son). Except for hypertension in the patient's brother and mother no other cardiovascular disease was present within the family (including coronary artery disease, stroke, dilated/hypertrophic/restrictive/arrhythmogenic cardiomyopathy, other forms of arrhythmias, or sudden cardiac death and no other family member had or has an implanted cardioverter defibrillator/pacemaker. The patient himself had no pacemaker or cardioverter/defibrillator implanted (Patient's characteristics are summarized in Table S1). In March 2016, at the patient's age of 66, blood samples were obtained for iPSC reprogramming purposes during a regular clinical follow-up visit.

#### Patient II (SHOX2 c.\*28T>C)

The male harboring the SHOX2 c.\*28T>C mutation was born in 1970 and developed AF at the age of 37 years. At the time of study enrollment in early 2011, his ECG showed a prolonged PR interval (first-degree atrioventricular block), a clinically unremarkable QRS duration (no bundle branch block) and normal QT durations. He was further diagnosed with DCM (left ventricular end diastolic diameter of 67 mm in February 2011). For primary prevention purposes, a single chamber cardioverter-defibrillator had been implanted. The cardiovascular risk factors of this patient included: Diabetes mellitus type II, hypercholesterolemia, ex-smoker status. Additionally, the patient suffered from chronic kidney disease and Crohn's disease. No other family member developed AF (parents, six brothers and sisters of the patient's father, six brothers and sisters of the patient, three children of the patient). The patient's father who had a known dilated cardiomyopathy died from myocardial infarction at the age of 46 and one of the father's brothers suffers from coronary artery disease and received surgical revascularization (coronary artery bypass). Another brother of the patient's father died from cardiac arrest at the age of 59. Several relatives (both parents, one brother of the patient's father, two brothers and one sister of the patient) suffer from hypertension. Besides that, no other cardiovascular disease was present within the family (including stroke, hypertrophic/restrictive/arrhythmogenic cardiomyopathy, or other forms of arrhythmias) and no other family member had or has an implanted cardioverter defibrillator/pacemaker. In March 2011, the patient suffered from progressive congestive heart failure with an ejection fraction (EF) of 25% presumably due to the patient's DCM and the AF that progressed to a persistent type. Consequently, the patient received a heart transplantation in October 2011. At that time, the patient was treated with Amiodarone (class III anti-arrhythmic drug) and Dobutamine (emergency medication for decompensated heart failure) besides the standard treatment with diuretics while β-blockers (class II anti-arrhythmic drug) and ACE inhibitors were stopped. Patient's characteristics are summarized in Table S1. In January 2015 at the patient's age of 45, blood samples were obtained for iPSC reprogramming purposes during a routine follow-up visit after heart transplantation.

		AF patient I ( <i>SHOX2</i> c.849C>A)	AF patient II ( <i>SHOX2</i> c.*28T>C)
	age at disease onset [years]	56	37
	age at recruitment [years]	66	41
	AF disease progression (at time of enrollment)	persistent	persistent
	heart transplantation	no	yes
Echocardiography	ejection fraction	43%	25%
(at time of enrollment)	structural	LA diameter = 43 mm	LVEED = 67  mm
	frequency [beats/min]	59	71
<b>F</b> 14	P wave [ms]	80	110
Electro- cardiogram	PR interval [ms]	180	210
caraiogram	QRS interval [ms]	70	90
	QT interval (QTc) [ms]	420 (416)	420 (457)
	medication	Pantozol, Marcumar	Amiodarone, Dobutamine, β-blockers (paused), ACE inhibitors (paused)
	cardioverter-defibrillator [yes/no]	Yes	No
	additional cardiovascular diseases	arterial hypertension	dilated cardiomyopathy
	cardiovascular risk factors	N/A	Diabetes mellitus II, hypercholesterolemia ex-smoker
	secondary diagnosis	N/A	Renal insufficiency, Morbus Crohn

Table S1. Patient characteristics. Related to Figure 1

Sample Name	Cell line	DNA quality	Best hit with database	Best Hit %	Present in database?	Cross- contamination	Identity confirmed	Genotype code
Control_1	hiPSCs	ok	MSTO-211H	90	no	no	unique sequence	ATTTATAAATATAAAAATATAATTT TAAAAAAATATATTTTTT
SHOX2_c.849C>A_1	hiPSCs	ok	WILL-1	90	no	no	unique sequence	ATTTATATAAAATAAAATTAAAATTT TAAAAAAATAAAAAA
SHOX2_c.849C>A_2	hiPSCs	ok	WILL-2	90	no	no	unique sequence	ATTTATATAAAATAAAATTAAAATTT TAAAAAAATAAAAAA
SHOX2_c.*28T>C_1	hiPSCs	ok	BSM	88	no	no	unique sequence	TTTTATATATAAAATATTTTTAATTTT ATATATATAAAAATTAATTTTAA
SHOX2_c.*28T>C_2	hiPSCs	ok	BSM	88	no	no	unique sequence	TTTTATATATAAAATATTTTTAATTTT ATATATATAAAAATTAATTTTAA
SHOX2_c.*28T>C_3	hiPSCs	ok	BSM	88	no	no	unique sequence	TTTTATATATAAAATATTTTTAATTTT ATATATATAAAAATTAATTTTAA
SHOX2_c.849C>A_isoWT	hiPSCs	ok	NK-92	92	no	no	unique sequence	ATTTATATAAAATAAAATTAAAATTT TAAAAAAATAAAAAWTTATTTAT
SHOX2_c.*28T>C_isoWT	hiPSCs	ok	BSM	88	no	no	unique sequence	TTTTATATATAAAATATTTTTAATTTT ATATATATAWAATTAATT

## Table S2. Multiplex cell line authentication of hiPSC lines. Related to Figure 1+5

Multiplex Human Cell Line Authentication Testing confirmed a shared origin for patient-derived *SHOX2* c.849C>A and *SHOX2* c.\*28T>C clones, as well as the common origin of isogenic controls with their respective patient lines. Cell lines were authenticated using Multiplex Cell Authentication (MCA) by Multiplexion (Heidelberg, Germany) as previously described (Castro et al., 2013). hiPSC lines were not cross-contaminated with other lines and were not present in cell line data bases.

## [Excel file 1]

## Table S3 Sequences of allele-quantification in sib-selection. Related to Figure 3

Detailed sequences and read numbers for each sib-selection analyzed via NGS are listed here. For allele quantification, only reads with a frequency higher than 0.25% of total reads (minimum allele frequency in 200 cells) were considered. These reads were classified into three categories: wildtype ('WT', no detectable *SHOX2* c.849C>A or *SHOX2* c.\*28T>C mutations), mutant ('Mut', detectable *SHOX2* c.849C>A or *SHOX2* c.\*28T>C mutations) or non-assignable ('N/A', where position *SHOX2* c.849C c.\*28\* were deleted).

#### SHOX2 c.849C>A gRNA-1

	Coordinates	Strand	MM	target_sequence	Distan	ce	Gene name	Gene ID	Seq?
1	chr10:78752841-78752863	-	4	G <mark>GGA</mark> CAGC[TGG <mark>C</mark> CAGCGCTG]	2179	Ι	KCNMA1	ENSG00000156113	ok
2	chr19:456297-456319	+	4	GCTCCTGG[GGGACAGCGCTG]	3052	Ι	RNA5SP462	ENSG00000252539	ok
3	chr21:46720679-46720701	+	4	GCAGGCGC[TGCACAGCGCTG]	225	Ι	LINC00315	ENSG00000184274	ok
4	chr8:110657693-110657715	-	4	GCAGCTGC[AGGGCAGCGCTG]	0	Е	RP11-422N16.3	ENSG00000248050	ok
5	chr2:26256570-26256592	-	4	TCTCCCGC[TGGACTGCGCTG]	384	-	RAB10	ENSG0000084733	ok
6	chr20:31573388-31573410	-	4	GCTGGAAG[TGGCCAGCGCTG]	132	Ι	SUN5	ENSG00000167098	ok
7	chr2:242667205-242667227	+	4	GCTACCCC[TGGAAAGCGCTG]	0	Е	ING5	ENSG00000168395	ok
8	chr1:151777738-151777760	+	4	ACTGCATC[TGTCCAGCGCTG]	0	Е	LINGO4	ENSG00000213171	ok
9	chr20:24597308-24597330	+	3	GAGGCAGC[TGGACAGGGCTG]	19519	Ι	SYNDIG1	ENSG00000101463	ok
10	chr9:137772667-137772689	+	3	GATGGAGC[TGGACAGAGCTG]	0	Е	FCN2	ENSG00000160339	ok
11	chr14:92673870-92673892	+	4	CCGGCAGC[TGGGGAGCGCTG]	43115	-	CPSF2	ENSG00000165934	ok
12	chr2:86260801-86260823	-	4	GCGGGATC[TGGACAACGCTG]	0	Е	POLR1A	ENSG0000068654	ok

#### SHOX2 c.\*28T>C gRNA-1

	Coordinates	Strand	MM	target_sequence	Distance	Gene name	Gene ID	Seq?
1	chr19:31785306-31785328	-	4	AAGTGCTG[ACTGAGTGCCGC]	11066 <mark>I</mark>	TSHZ3	ENSG00000121297	ok
2	chr6:30554074-30554096	+	4	GTCTGCAT[GGTGAGTGCCGC]	0 E	ABCF1	ENSG00000204574	ok
3	chr11:118793326-118793348	-	3	GGGTACAG[GCTGGGTGCCGC]	2525 <mark>I</mark>	UPK2	ENSG00000110375	ok
4	chr22:41809060-41809082	+	3	GCGCGCCG[GCTGTGTGCCGC]	13730 -	TEF	ENSG00000167074	ok
5	chr3:32863249-32863271	+	4	GCG <mark>G</mark> GC <mark>CC</mark> [GCTG <mark>C</mark> GTGCCGC]	2825 <mark>I</mark>	TRIM71	ENSG00000206557	ok
6	chr14:105995090-105995112	+	4	GCCCGCTG[GCTGAGCGCCGC]	0 E	TMEM121	ENSG00000184986	ok
7	chr7:55433418-55433440	+	3	GCGAGCAG[GCTGTGAGCCGC]	0 E	LANCL2	ENSG00000132434	ok
8	chr14:55031837-55031859	-	4	GCG <mark>CAG</mark> AG[GCTGAGTG <mark>G</mark> CGC]	1956 -	SAMD4A	ENSG0000020577	ok
9	chr8:103646515-103646537	-	4	AGATGCAG[GCTGAGTGCGGC]	11956 -	POU5F1P2	ENSG00000253382	ok
10	chr1:2383625-2383647	+	4	CGGTGGAG[GCTGAGTGCAGC]	15251 I	PLCH2	ENSG00000149527	ok
11	chr17:39765137-39765159	+	4	GACTGAAG[GCTGAGTGCTGC]	871 -	KRT16	ENSG00000186832	ok

#### Table S4. Off-target analysis for SHOX2 c.849C>A gRNA-1 and SHOX2 c.\*28T>C gRNA-1. Related to Figure 4

The top 20 off-target sites predicted by *CCTop* were further evaluated. Exonic off-target sites were automatically included into downstream analysis. For intronic and intergenic off-target sites, the target sequence coordinates were analyzed in the *UCSC Genome Browser* (https://genome.ucsc.edu/). The inclusion criteria for downstream analysis were: Conservation among species, DNase clustering, expressed sequence tags (ESTs) and active chromatin marks. If a combination of these criteria indicated a potential regulatory relevance of this DNA segment, the off-target site was sequenced. In total, 12 off-targets were analyzed per isogenic clone. Nucleotides in **red** are mismatches, MM = mismatches, I = intronic, I = intergenic

## [Excel file 2]

## Table S5. Primers for *in vitro* synthesis, dPCR, NGS, qRT-PCR, Sanger Sequencing, and sequences of single-strand oligodeoxynucleotides. Related to Figure 1-6

All primer sequences used for *in vitro* synthesis, dPCR, NGS, Sanger Sequencing, qRT-PCR of pluripotency and cardiomyocyte subtype marker genes are listed here. Additionally, the sequences of single-stranded oligodeoxynucleotides that were transfected together with Cas9 RNP/gRNA complexes to serve as a repair template are given.

Target	Host	Source	Catalog number	Dilution
anti-rabbit IgG-AlexaFluor488	goat	Thermo Fisher Scientific	A32731	1:500
anti-mouse IgG-AlexaFluor488	goat	Thermo Fisher Scientific	A11029	1:500
anti-rabbit IgG-AlexaFluor594	goat	Thermo Fisher Scientific	A32740	1:500
anti-rabbit IgG-AlexaFluor647	goat	Thermo Fisher Scientific	A27040	1:500
NANOG	rabbit	Abcam	ab21624	1:500
OCT4	rabbit	Abcam	ab19857	1:500
SOX2	rabbit	Abcam	ab137385	1:200
TRA1-81-AlexaFluor488	mouse	BD Pharmingen	560174	1:20
Cardiac Troponin T (EPR3696)	rabbit	Abcam	ab45932	1:500
$\alpha$ -actinin (sarcomeric; clone EA-53)	mouse	Sigma	A7811	1:300

 Table S6. Primary and secondary antibodies. Related to Figure 1+5+6

## **Supplemental Experimental Procedures**

## **Ethics statement**

The study was approved by the ethical commission of the Medical Faculty, Technical University of Munich, Munich, Germany ("Induzierbare pluripotente Stammzellen als innovatives Patienten-basiertes *in vitro* Modell für Vorhofflimmern" als Teilprojekt 5 des Hauptprojekts "Erzeugung und Charakterisierung patientenspezifischer induzierter, pluripotenter Stammzellen" 2109/08) and was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. Every participant gave written informed consent including the consent to use their blood samples prior to the inclusion in the study.

## hiPSC Generation

AF patients carrying the *SHOX2* mutations were recruited from the Department of MedicineI of the Ludwig-Maximilians-University Hospital Grosshadern, Munich. Peripheral blood mononuclear cells were isolated and reprogrammed into hiPSCs using CytoTune-iPS 2.0 Sendai Reprogramming Kit (Life Technologies) as previously described (Moretti et al., 2010). Loss of Sendai virus in hiPSCs was confirmed at passage 10-20 by RT-PCR of the Sendai vector and viral transgenes OCT4, SOX2, KLF4 and c-MYC. Pluripotency of hiPSCs was verifed by alkaline phosphatase staining, qRT-PCR analysis of pluripotency markers *OCT4*, *SOX2*, *NANOG*, *REX1* and *TDGF-1* and immunofluorescence analysis for NANOG, TRA-1–81, OCT4 and SOX2. Spontaneous differentiation of hiPSCs into cells of all three germ layers was induced by embryoid body formation, as previously reported (Moretti et al., 2010). Expression of lineage markers specific for endoderm (*PDX1*, *SOX7*, *AFP*), mesoderm (*CD31*, *DES*, *ACTA2*, *MYL2*, *SCL*, *CDH5*) ectoderm (*KRT14*, *NCAM1*, *TH*, *GABRR2*) was assessed at day 21 of EB differentiation.

## Immunohistological analysis

Four days after plating, cardiomyocytes were fixed (4% PFA for 15 min at RT), permeabilized and blocked with 10% goat serum in PBS/0.1% Triton-X-100 for one hour1h at 37°C. Cells were stained with primary antibodies cTNT and  $\alpha$ -actinin in PBS/0.1% Triton-X-100 containing 1% goat serum overnight at 4°C. AlexaFluor488- and AlexaFluor-594-conjugated sSecondary antibodies (Thermo Fisher Scientific) specific to the appropriate species were used in PBS/0.1% Triton-X-100 containing 1% goat serum for one hour1h at 37°C. Nuclei were detected with 1 µg/ml Hoechst 33258. A list of the antibodies is provided in Table S6. Images were acquired with a SP8 Leica confocal microscope and processed with Photoshop.

## qRT-PCR

For the analysis of germ layer and pluripotency markers, total RNA was extracted using the Absolutely RNA Miniprep Kit (Agilent Technologies) according to manufacturer's recommendations. For the gene expression analysis of cardiomyocyte markers, total RNA was extracted with guanidinium thiocyanate ('Trizol'). cDNA reverse transcription was performed using the SuperScript III First-Strand Synthesis Kit (Invitrogen). For semi-quantitative analysis, 1  $\mu$ l cDNA was subjected to PCR reaction using Taq polymerase (Thermofisher Scientific). All samples were measured in duplicates and the relative gene expression levels were normalized to the reference gene *GAPDH* semi-quantitively. For qRT-PCR analysis, measurements were conducted on a QuantStudio3 System (Thermo Fisher Scientific) using 2  $\mu$ l cDNA and the qPCRBIO SyGreen Mix (Nippon Genetics). All samples were measured in duplicates and the relative gene expression levels were normalized to the reference gene *SDHA* and *HPRT1* using the Relative Standard Curve. A list of the primers is provided in Supplementary Table 10.

### Alkaline phosphatase activity screening

Alkaline phosphatase activity was analyzed using the NBT/BCIP alkaline phosphatase blue substrate (Roche), according to the manufacturer's instructions. Images were acquired with a DMI6000-AF6000 Leica microscope and processed with ImageJ or Photoshop.

### Karyotyping

Chromosomes were obtained according to routine procedures and based on previously published protols. Karyograms were made of trypsin giemsa stained Metaphases. For multiplex fluorescence in situ hybridization (M-FISH) of the hiPSC lines, cells were arrested in methaphase with N-desacetyl-N-methylocolchicine (Thermo Fisher Scientific) and chromosomes were prepared with standard methods. Representative images (n = 10) were recorded using a DM-RXA epifluorescence microscope and processed using the Leica MCK software.

### Sanger Sequencing of annotated homozygous WT clones and off-targets

The Cas9 target and potential off-target regions with several hundred nucleotides up- and downstream were amplified using the HotStarTaq Polymerase (QIAGEN) according to manufacturer's instructions. PCR products were sent for Sanger Sequencing to GENEWIZ (Germany) and screened for mutations with the SnapGene software (GSL Biotech).

#### Flow cytometry analysis

For differentiation efficiency analysis via flow cytometry, d20 hiPSC-derived cardiomyocytes were fixed with 4% PFA for 15 min at RT, blocked with 10% goat serum in PBS/0.1% Triton-X-100 for 1h at RT and incubated with cTNT antibody (1:500; abcam) at 4°C overnight. Subsequently, anti-rabbit AlexaFluor-647-conjugated secondary antibody was used (1:500; Life Technologies) for 1h at RT. Flow cytometry acquisition was performed on Gallios (Beckman coulter), and data were analyzed using Kaluza software (Beckman coulter).

#### Statistical analysis

For gene expression analysis, data are presented as mean  $\pm$  SD. GraphPad Prism 7 was used to perform multiple t-tests followed by the Holm-Sidak post-hoc test for multiple comparisons.

## **Supplemental References**

CASTRO, F., DIRKS, W. G., FAHNRICH, S., HOTZ-WAGENBLATT, A., PAWLITA, M. & SCHMITT, M. 2013. High-throughput SNP-based authentication of human cell lines. *Int J Cancer*, 132, 308-14.

MORETTI, A., BELLIN, M., WELLING, A., JUNG, C. B., LAM, J. T., BOTT-FLUGEL, L., DORN, T., GOEDEL, A., HOHNKE, C., HOFMANN, F., SEYFARTH, M., SINNECKER, D., SCHOMIG, A. & LAUGWITZ, K. L. 2010. Patient-specific induced pluripotent stem-cell models for long-QT syndrome. *N Engl J Med*, 363, 1397-409.