Stem Cell Reports
Resource
ISSCR
OPEN ACCESS

Precise Correction of Heterozygous *SHOX2* Mutations in hiPSCs Derived from Patients with Atrial Fibrillation via Genome Editing and Sib Selection

Simon Alexander Sumer,[1,2,7] Sandra Hoffmann,[1,2,7] Svenja Laue,[3,4] Birgit Campbell,[3] Kristin Raedecke,[1,2] Viktoria Frajs,[1] Sebastian Clauss,[4,5] Stefan Kääb,[4,5] Johannes W.G. Janssen,[6] Anna Jauch,[6] Karl-Ludwig Laugwitz,[3,4] Tatjana Dorn,[3] Alessandra Moretti,[3,4,8] and Gudrun A. Rappold[1,2,8,*]

[1]Department of Human Molecular Genetics, Institute of Human Genetics, University of Heidelberg, 69120 Heidelberg, Baden-Wuerttemberg, Germany
[2]DZHK (German Center for Cardiovascular Research), Partner Site Heidelberg/Mannheim, Heidelberg, Germany
[3]First Department of Medicine, Cardiology, Klinikum Rechts der Isar – Technical University of Munich, 81675 Munich, Bavaria, Germany
[4]DZHK (German Center for Cardiovascular Research), Partner Site Munich, Munich, Germany
[5]Department of Medicine I, Klinikum Grosshadern, University of Munich (LMU), 81675 Munich, Bavaria, Germany
[6]Department of Human Genetics, Institute of Human Genetics, University of Heidelberg, 69120 Heidelberg, Baden-Wuerttemberg, Germany
[7]These authors contributed equally
[8]Co-senior author
*Correspondence: gudrun.rappold@med.uni-heidelberg.de
https://doi.org/10.1016/j.stemcr.2020.08.015

SUMMARY

Patient-specific human induced pluripotent stem cells (hiPSCs) offer unprecedented opportunities for the investigation of multigenic disease, personalized medicine, and stem cell therapy. For heterogeneous diseases such as atrial fibrillation (AF), however, precise correction of the associated mutation is crucial. Here, we generated and corrected hiPSC lines from two AF patients carrying different heterozygous *SHOX2* mutations. We developed a strategy for the scarless correction of heterozygous mutations, based on stochastic enrichment by sib selection, followed by allele quantification via digital PCR and next-generation sequencing to detect isogenic subpopulations. This allowed enriching edited cells 8- to 20-fold. The method does not require antibiotic selection or cell sorting and can be easily combined with base-and-prime editing approaches. Our strategy helps to overcome low efficiencies of homology-dependent repair in hiPSCs and facilitates the generation of isogenic control lines that represent the gold standard for modeling complex diseases *in vitro*.

INTRODUCTION

Research on the genetic basis of human cardiovascular disease has mainly relied on animal models in the past due to the limited access to human primary cardiac tissue. However, species-specific electrophysiological and transcriptional differences in cardiomyocytes can confound the translation of these findings to clinical relevance (Brandao et al., 2017; Moretti et al., 2013). Human induced pluripotent stem cells (hiPSCs) can help to overcome these limitations, as they provide an unlimited source of cardiomyocytes for the investigation of pathogenic mechanisms and the discovery of promising therapeutic targets (Abou-Saleh et al., 2018). Patient-derived hiPSCs represent a personalized drug-screening platform for precision medicine and individualized therapy (van Mil et al., 2018), Yet, the phenotype of the derived cardiomyocytes remains largely immature and the complex three-dimensional tissue structure can be mimicked only insufficiently *in vitro* (Moretti et al., 2013). In addition, substantial interline variability in cardiac differentiation capacity exists, emphasizing the need for reliable controls. Nevertheless, hiPSCs have been successfully used to model several cardiovascular diseases (Brandao et al., 2017; El-Battrawy et al., 2018; Goedel et al., 2017; Moretti et al., 2020).

The short stature homeobox 2 (SHOX2) transcription factor plays a crucial role in the development of the sinoatrial node, the native pacemaker of the heart. By antagonizing *NKX2.5*, it prevents the formation of working myocardial tissue, simultaneously activating the gene program of pacemaker cells, such as *ISL1* and *HCN4* (Blaschke et al., 2007; Hoffmann et al., 2013; Ye et al., 2015). We and others have shown that mutations in this gene are associated with early-onset and familiar forms of atrial fibrillation (AF) and sinus node dysfunction (Hoffmann et al., 2016, 2019; Li et al., 2018). In a previous screening of patients with an early-onset form of AF, we identified two mutations that interfere with the function of *SHOX2*: the coding mutation *SHOX2* c.849C>A (= SHOX2 p.H283Q) was shown to impede the protein's function as a transcriptional activator, as demonstrated by the failed rescue of the bradycardia phenotype in a zebrafish model after knockdown of endogenous Shox2. Another *SHOX2* variant resides in the 3′ UTR (*SHOX2* c.*28T>C), leading to a novel microRNA binding site for *miR-92b-5p*, which has been shown to be functional *in vitro*. AF patients harboring this variant have reduced microRNA expression levels in the plasma compared with AF-affected non-carriers and significantly prolonged PR intervals (Hoffmann et al., 2016). To develop an *in vitro* AF model, we recruited two patients with these *SHOX2* mutations (*SHOX2* c.849C>A

or *SHOX2* c.*28T>C) for reprogramming of somatic cells to hiPSCs.

Recent advances in gene transduction and editing technologies have potentiated the applicability and versatility of hiPSCs (Brookhouser et al., 2017). Genome targeting has become feasible due to the development of site-specific nucleases, such as clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9. Following a Cas9-induced double-strand break (DSB), endogenous repair mechanisms are activated by the cell: the error-prone non-homologous end joining (NHEJ) or the precise homology-directed repair (HDR), which uses (partially) homologous DNA to restore the DNA sequence (Jackson and Bartek, 2009). In the presence of an exogenously introduced template, the alteration or insertion of specific DNA sequences, ranging from single-nucleotide exchanges to whole transgenes, can be achieved. Especially, the correction of disease-associated mutations in patient-specific hiPSCs holds great potential for precision disease modeling (Jang and Ye, 2016). These so-called isogenic controls differ only in single genetic variants, allowing for a direct genotype-to-phenotype comparison. This is of particular interest when studying multigenic diseases such as AF, where combinations of different variants with low impact are contributing to the observed phenotype.

A key limitation in the genome editing of hiPSCs is the fact that these cells preferentially choose the error-prone NHEJ pathway rather than the precise homology-directed approach (Guo et al., 2018). Under normal circumstances, frequencies of ~1% for single-base substitutions or deletions are often reported (Miyaoka et al., 2014; Soldner et al., 2011). In recent years, substantial efforts have been made to develop strategies improving HDR efficiency, ranging from changes in the culturing conditions to suppressing NHEJ via small molecules or gene knockdown, modifications of the Cas9 enzyme, and optimization of the donor template (Anzalone et al., 2019; Chu et al., 2015; Guo et al., 2018; Komor et al., 2016; Okamoto et al., 2019; Yu et al., 2015).

Here, we propose a strategy to precisely correct heterozygous mutations adapted from a previously published method (Miyaoka et al., 2014), which was developed for the insertion of single-base substitutions into a wild-type (WT) background. It is based on the stochastic enrichment of precisely edited cells by subdivision of genome-edited cells into small pools and the detection of these subpopulations via allele quantification. With this approach, we were able to enrich our cells of interest 8- to 20-fold compared with the initial editing efficiency, allowing for a less work-intense screening for rare editing events, which can be combined with other optimization approaches. The corrected hiPSCs will help to elucidate the function of *SHOX2* in the genetic network of atrial and nodal cardiomyocytes and its contribution to the development and progression of AF.

RESULTS

Sib Selection Stochastically Enriches Isogenic Subpopulations That Are Detected by Allele Quantification

We generated hiPSCs derived from peripheral blood mononuclear cells of two AF patients harboring heterozygous *SHOX2* mutations (*SHOX2* c.849C>A and *SHOX2* c.*28T>C) using non-integrating Sendai viruses encoding KLF4, OCT3/4, SOX2, or c-MYC (Figure S1A) (for detailed patient characteristics see Table S1). The presence of the patient-specific *SHOX2* c.849C>A and *SHOX2* c.*28T>C mutations was confirmed by Sanger sequencing (Figure S1B). All derived clones (two clones from patient I and three clones from patient II) displayed stem cell-like features and pluripotency capacity, as shown by high activities of alkaline phosphatase (Figure S1C), had lost the viral transgenes after 10–20 passages (Figure S1D), expressed pluripotency markers on protein and RNA level (Figures S1E and S1F), and differentiated spontaneously into all three germ layers (Figure S1G). Moreover, they were karyotypically normal (Figure S1H) and shared their respective patient origin (Table S2). Clones no. 1 from patient I and patient II were used for gene correction.

To target the heterozygous mutations, two single guide RNAs (gRNAs) per locus with cutting sites close to the mutations were selected (Figure 1). The gRNAs were predicted to range from moderately to highly efficient, with a substantial number of off-targets. However, as gene conversion tracks are relatively short in mammalian cells (Elliott et al., 1998), the distance between the Cas9 cut site and the targeted DNA sequence has to be minimized to achieve high HDR efficiency. Cells were transfected with Cas9 ribonucleoprotein (RNP)/gRNA complexes and a single-stranded oligodesoxynucleotide (ssODN) as the HDR template (Figure 1A). After 48 h the genome-targeting efficiency of each gRNA was determined by next-generation sequencing (NGS) for a precise estimation of indel size, frequency, and sequence identity. For the *SHOX2* c.849C>A locus, the two selected gRNAs were moderately effective, producing indel frequencies of 36% and 21%, respectively. For the *SHOX2* c.*28T>C locus, gRNA-1 and gRNA-2 were less effective, producing indel frequencies of 17% in hiPSCs (Figure 1B). Due to the heterozygous nature of the *SHOX2* mutations, the frequency of HDR events could not be precisely determined in these large cell pools.

The fractionation of a heterogeneous population of genome-edited stem cells can randomly enrich desired subpopulations, such as isogenic cells. A population of cells

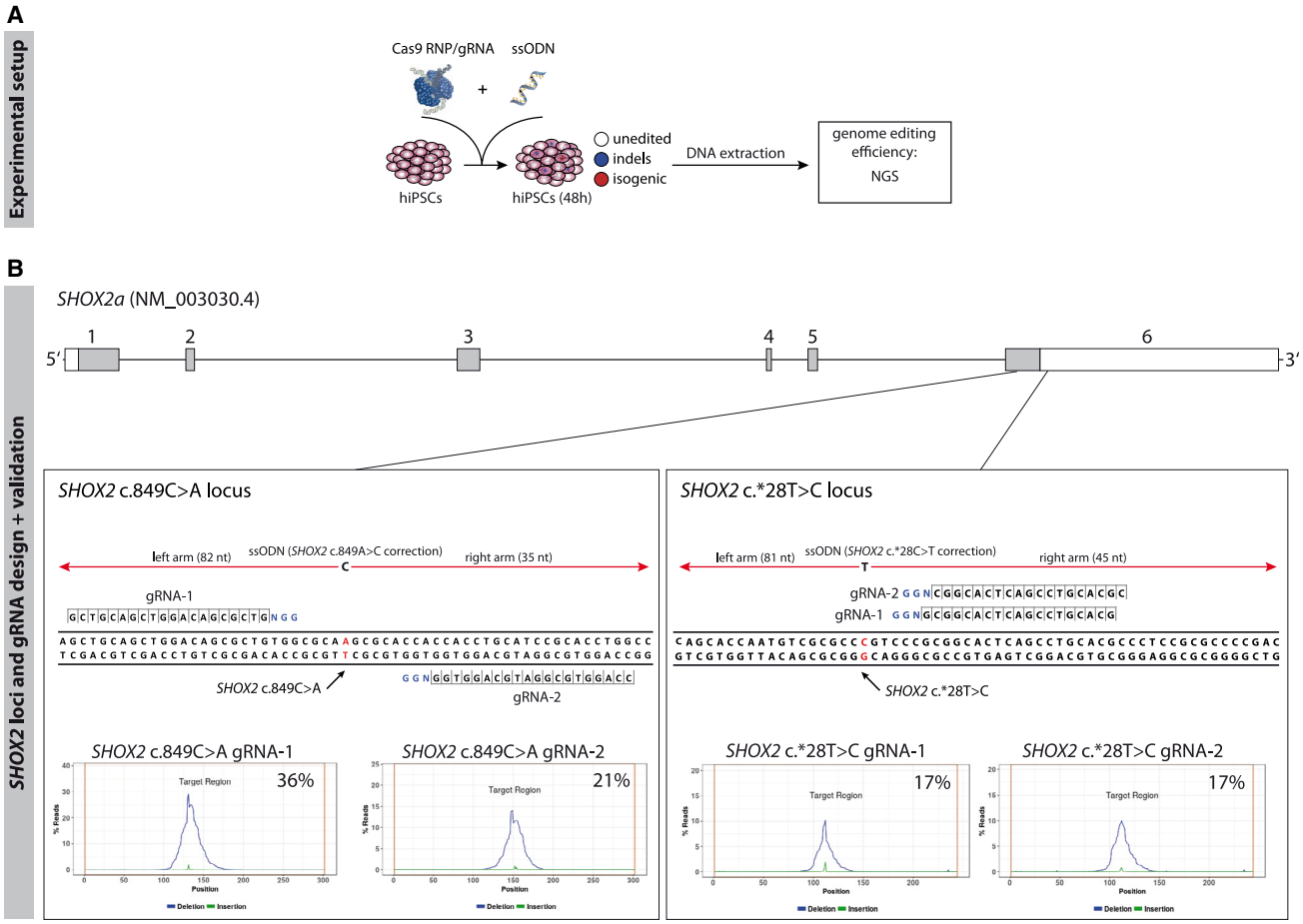


Figure 1. gRNA Design and Validation for the *SHOX2* c.849C>A and the *SHOX2* c.*28T>C Loci

(A) Experimental overview of gRNA validation. hiPSCs were transfected with Cas9 RNP/gRNA complexes and ssODNs. After 48 h, the Cas9-targeted region was amplified and analyzed via NGS.

(B) Boxes, top: The *SHOX2* c.849C>A and *SHOX2* c.*28T>C loci with selected gRNA binding sites. PAM sequences are depicted in blue, the mutated base pair is in red. Boxes, bottom: Indel frequency for each gRNA in hiPSCs 48 h after transfection. The targeted region is centered. The cumulative frequency of deletions (blue) and insertions (green) at each position is depicted in percentage of reads (% Reads). Abbreviations: RNP, ribonucleoprotein; gRNA, single guide RNA; ssODN, single-stranded oligodesoxynucleotide; NGS, next-generation sequencing.

See also Figure S1, Table S1, S2, S5, and S6.

containing a small number of cells of interest is subdivided into small pools ("sib selection"). Of these, the one with the highest percentage of target cells is selected and subjected to a new round of subdivision. After the enrichment of target cells to a reasonable amount, single-cell cloning can be performed to achieve a pure cell population. We hypothesized that—by quantifying the ratio of WT and Mut alleles in these sib selections—we would be able to detect cells that had precisely corrected the heterozygous mutation. A present subpopulation of isogenic cells contributes two WT alleles to the DNA pool, leading to a shift in the WT/Mut allele ratio. Sib selections with an overaverage abundance of WT alleles potentially contain enriched amounts of isogenic cells and can, therefore, be selected for further analysis and single-cell cloning.

We utilized digital PCR (dPCR) to precisely quantify and identify DNA molecules to determine WT/Mut allele ratios within individual sib selections (Figure 2A). Using plasmids containing either the mutated or the unmutated *SHOX2* gene, we confirmed the specificity of our TaqMan probes (Figure 2B). The sensitivity and specificity of the system were tested by mixing different ratios of genomic DNA from healthy donors with DNA from patient-specific hiPSCs and comparing detected allele ratios to actual ratios. The strong correlation between calculated and measured Mut alleles indicated a high precision of allelic

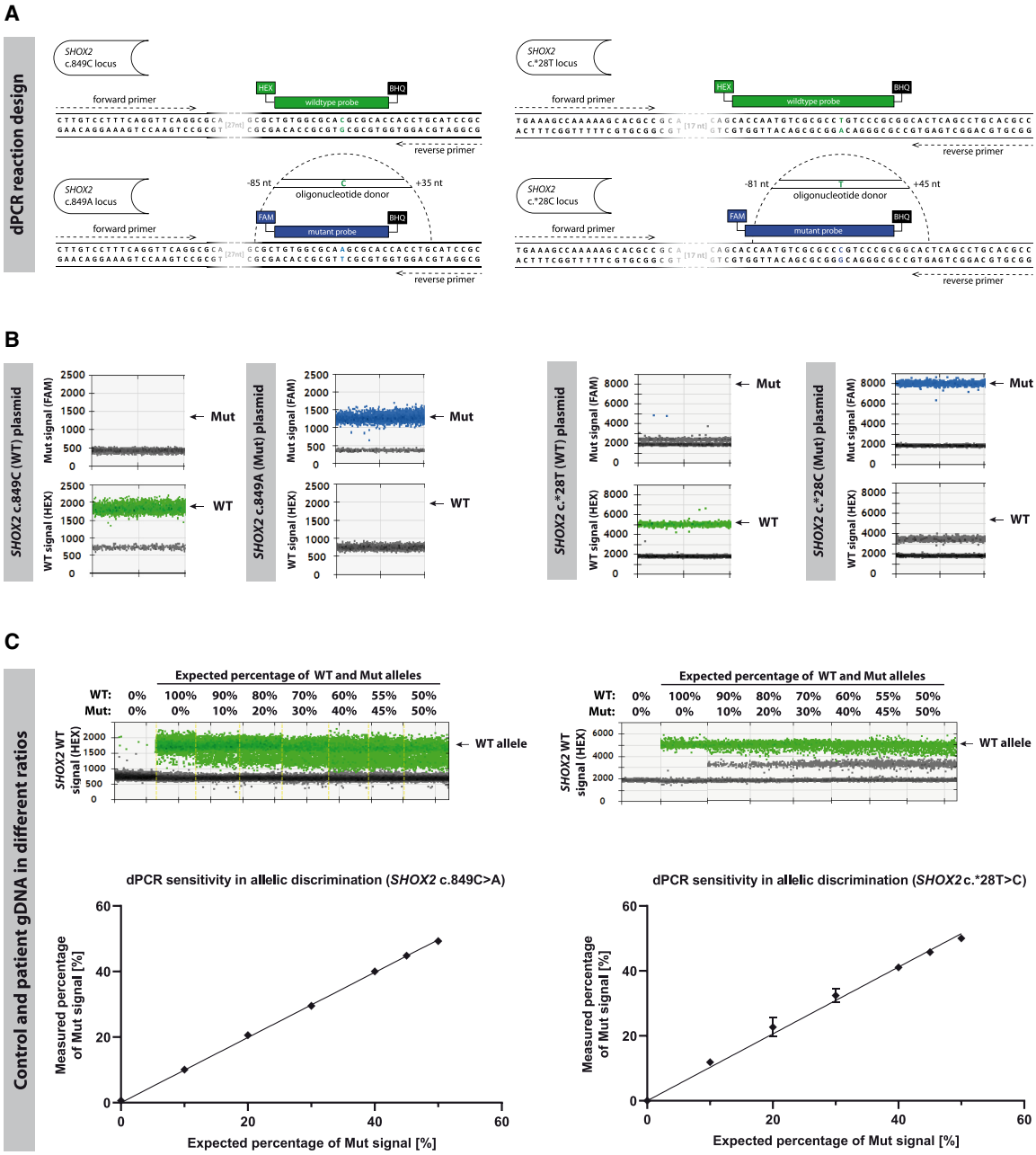


Figure 2. Pretests for Allele Quantification via dPCR

(A) Primer/probe design for detection of WT and Mut alleles with specific probes. The mutation-spanning oligonucleotide donor is depicted above the mutant allele.

(B) dPCR result for different ratios of control and patient genomic DNA represented as HEX channel 1D amplitude (top) and plotted against the expected percentage (bottom); n = 3, error bars represent ±SD of the mean.

(C) Probe specificity test with plasmids containing WT (SHOX2c.849C, SHOX2 c.*28T) and Mut (SHOX2 c.849A, SHOX2 c.*28C) alleles. Data are expressed as mean ± SD of three independent experiments. Abbreviations: dPCR, digital PCR; WT, wild type; Mut, mutant; gDNA, genomic DNA; FAM, 6-carboxyfluorescein; HEX, hexachloro-fluorescein; BHQ, Black Hole Quencher.

See also Table S5.

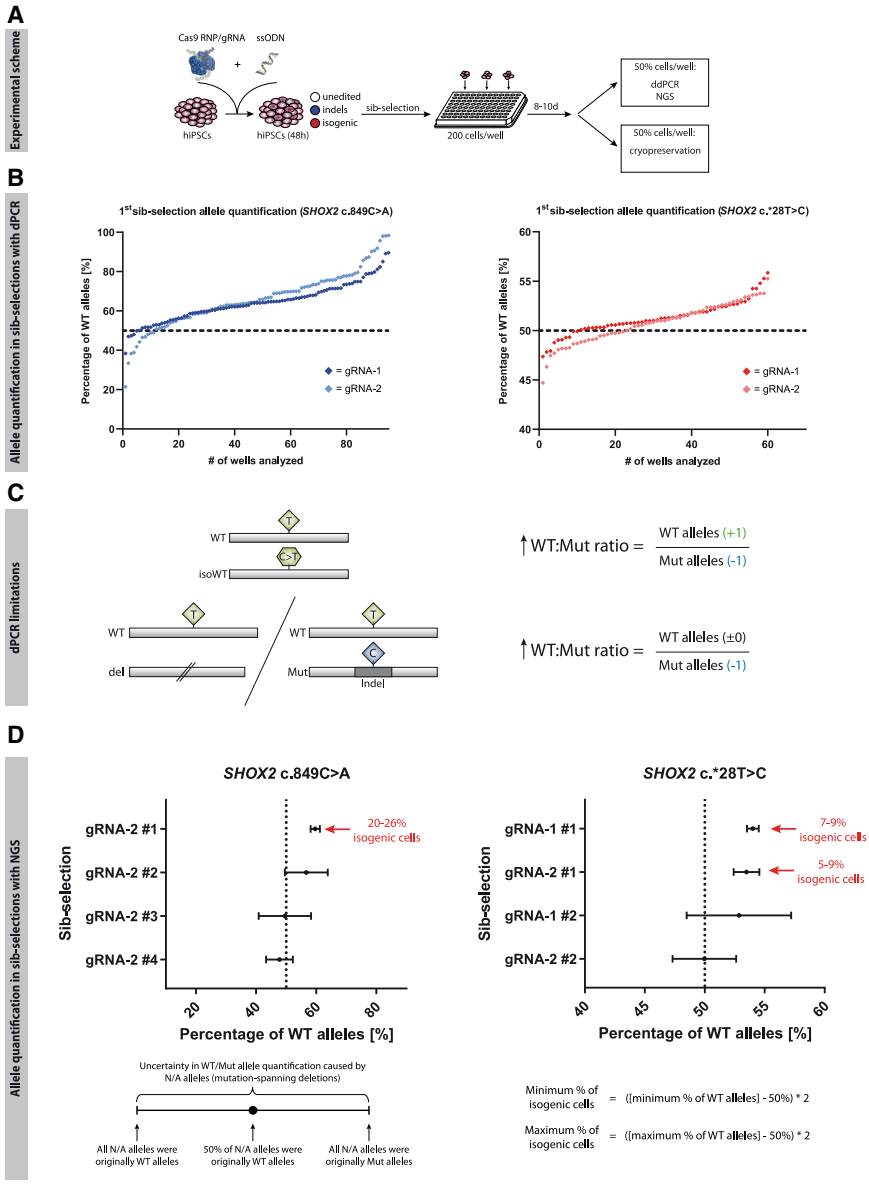


Figure 3. Allele Quantification in Sib Selections via dPCR and NGS

(A) Experimental scheme of gene-editing approach.

(B) WT and Mut alleles were quantified in each sib selection 10 days after transfection via dPCR. Each dot represents the result for one sib selection. The dotted line marks the 50% WT allele percentage expected in unedited hiPSCs. Sib selections with the highest abundance of WT alleles were thawed and re-analyzed.

(C) Limitations of dPCR in allele quantification: due to the PCR-based allele detection, non-amplifiable Mut alleles can lead to shifts in the WT/Mut ratio, similar to what is caused by isogenic subpopulations.

(D) Allele quantification via NGS: alleles with *SHOX2* c.849- and *SHOX2* c.*28-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. Sib selections in which an increased WT/Mut allele ratio is not solely explicable by a loss of detectable Mut alleles were chosen for single-cell cloning. The percentage of isogenic cells was calculated with the given formula. Abbreviations: Indel, insertion/deletion; N/A, non-assignable.

See also Figure S2, Table S5.

quantification as well as a specificity of WT and Mut TaqMan probes for their respective alleles (Figure 2C).

Cas9 RNP/gRNA/ssODN-transfected hiPSCs were seeded into small cell pools of 200 cells per well on a 96-well plate, 48 h after transfection and grown until confluency (~8–10 days) (Figure 3A). From each well, 50% of the cells were cryopreserved, while the other 50% were subjected to DNA extraction and subsequent allele quantification with dPCR (Figure 3B). Most of the samples after the sib selection still showed a nearly equal allele distribution of 50:50 (dotted line), which is characteristic of a heterozygous mutation and indicates no enrichment of targeted cells. However, a higher abundance of WT alleles was seen in some of the samples that were selected for further analysis.

Isogenic Subpopulations in Sib Selections Are Re-quantified via Next-Generation Sequencing

The major limitation of allele quantification via dPCR is its dependence on a functional PCR reaction. Mutations or the complete loss of primer/probe binding sites in alleles can prevent a successful amplification of the DNA strands and therefore their identification. If large populations of hiPSCs in a sib selection contain WT alleles and non-detectable mutant alleles, it will generally lead to a shift in the WT/Mut allele ratio similar to what is caused by isogenic cells (Figure 3C). Consequently, sequences had to be analyzed in detail via NGS to confirm the presence of isogenic cells in the chosen sib selections. A total of 11 sib selections were deep sequenced. To determine

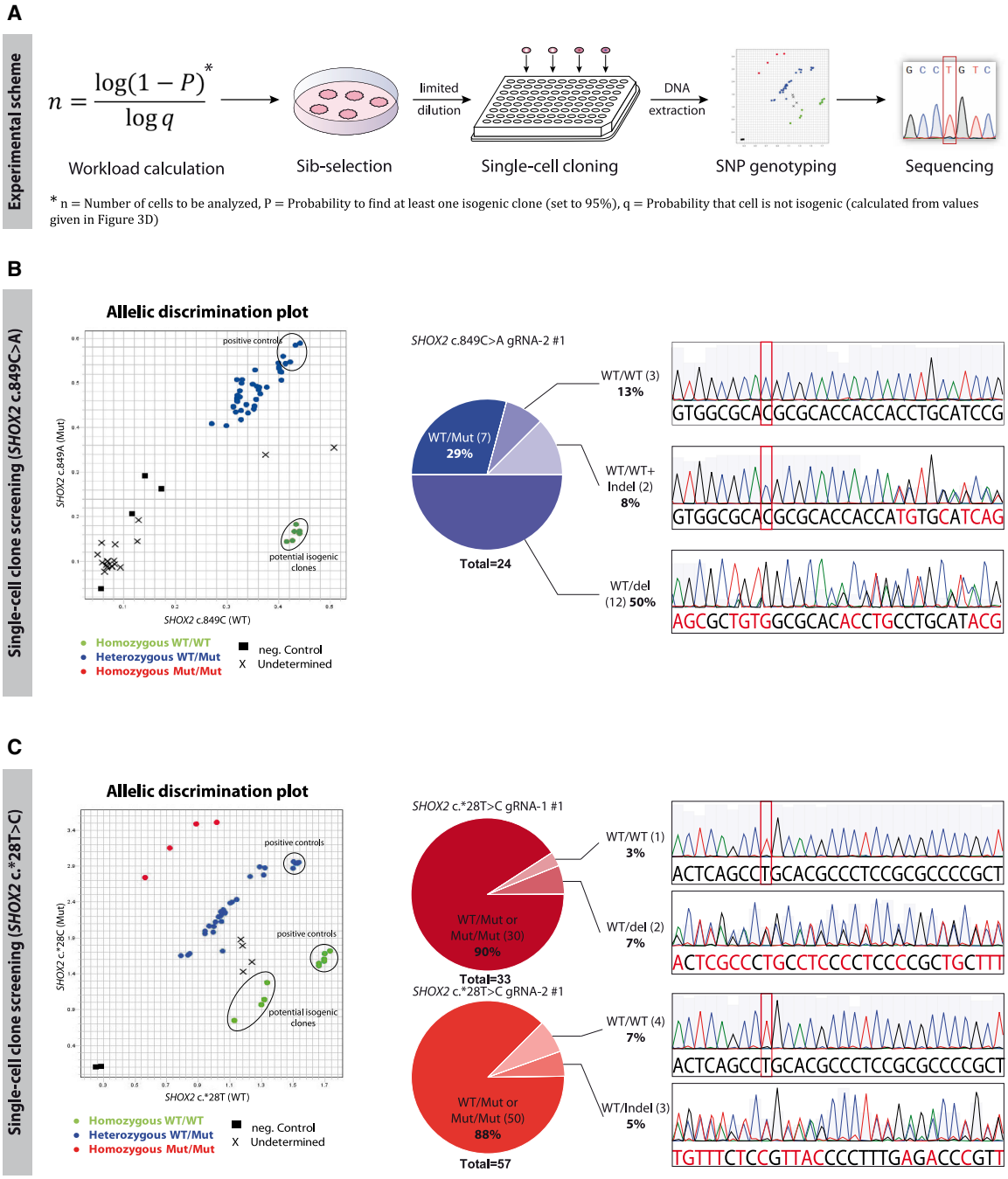


Figure 4. Single-Cell Cloning with Sib Selections and Screening for Isogenic Clones

(A) Experimental scheme for single-cell cloning and screening: the number of clones needed to be analyzed was calculated with a binomial distribution function. Sib selections with isogenic subpopulations were thawed for single-cell cloning via limited dilution. Clones were screened via TaqMan probe-based SNP genotyping, and potential homozygous WT clones were confirmed with Sanger sequencing.

(B) Screening for isogenic clones derived from heterozygous *SHOX2* c.849C>A cells: 24 single-cell clones were genotyped and sequenced. In 5/24 clones (21%) the mutation was corrected back to WT, with 3 clones showing no additional mutations several hundred nucleotides up- and downstream.

(legend continued on next page)

the percentage of isogenic cells from WT/Mut allele ratios, we made two assumptions: first, we considered the probability of *SHOX2* copy number variations (CNVs), for example, due to trisomy or gene duplications, to be very low. CNVs could lead to changes in allele ratios that are also not caused by isogenic subpopulations. Second, we neglected the probability of an imprecise correction of the *SHOX2* mutations (mutation lost, but other mutations introduced) for that moment. Under normal circumstances, HDR leads to a complete restoration of the sequence. This allowed us to directly calculate the percentage of isogenic subpopulations from the WT and Mut alleles (Figure 3D). We classified NGS reads into three categories, WT alleles, Mut alleles, and non-assignable (N/A) alleles, in which the base of interest (*SHOX2* c.849 or *SHOX2* c.*28) was deleted (for detailed sequences see Table S3). This classification was carried out independent of additional mutations in the respective alleles. N/A alleles caused uncertainty in allele quantification, as their origin from either WT or Mut alleles could not be determined. This uncertainty was addressed by defining all N/A alleles as either WT or Mut alleles when calculating the WT/Mut allele ratio. The result was a range of possible ratios, spanning the two extreme scenarios in which N/A alleles were counted as either all WT or all Mut (Figure 3D). Using gRNA-1 to correct the *SHOX2* c.849C>A mutation led to a high percentage of reads with deletions spanning the targeted mutation, which made the precise quantification of isogenic sib selections impossible (Figure S2). For the *SHOX2* c.849C>A sib selections gRNA-2 nos. 2–4, a large fraction of N/A alleles led to a wide range of potential allele ratios (Figure 3D). As this included a scenario in which the higher percentage of WT alleles could be explained solely by a loss of detectable Mut alleles, a subpopulation of isogenic cells was possible, but not guaranteed. On the other hand, sib selection *SHOX2* c.849C>A gRNA-2 no. 1 had a strong and robust increase in WT alleles, indicating a large fraction of isogenic clones. For *SHOX2* c.*28T>C sib selections, gRNA-1 no. 1 and gRNA-2 no. 1 showed similar shifts and were therefore selected for single-cell cloning together with *SHOX2* c.849C>A gRNA-2 no. 1. Allelic distributions were used to calculate the percentage of isogenic cells in these cell pools, estimating frequencies of 20%–26% for *SHOX2* c.849C>A gRNA-2 no. 1, as well as 7%–9% and 5%–9% in sib selection gRNA-1 no. 1 and gRNA-1 no. 2, respectively (Figure 3D). With a supposed initial efficiency of ~1% for precise genome editing, the sib-selection process led to a significant enrichment of target cells.

The number of cells that had to be screened to find at least one isogenic clone with a given probability was determined by negative binominal distribution. The applied parameters were the calculated frequency of target cells (~23% or ~8%), the desired number of positive clones to be found (≥1 clone), and a self-defined chance of success to find one (95%) (Figure 4A). Twelve clones for sib selection *SHOX2* c.849C>A gRNA-2 no. 1 and 35–40 clones for *SHOX2* c.*28T>C sib selections were generated to find at least one clone of interest with a 95% probability. Compared with the nearly 300 cells that would have to be screened for the same chances of success—if the target cell population was only ~1%—still, this represented a substantial reduction in screening workload.

Monoclonal cell populations were obtained via limiting dilution (Figure S3). Both dPCR primers and probes were reused for TaqMan SNP genotyping of *SHOX2* c.849C>A and of *SHOX2* c.*28T>C to preselect potentially homozygous WT clones. From the allele discrimination plot, all annotated homozygous WT clones were identified (Figures 4B and 4C). Subsequent sequencing confirmed the precise correction of the heterozygous *SHOX2* c.849C>A or *SHOX2* c.*28T>C mutation and was used to screen for additional mutations up- and downstream of the Cas9 target site. For *SHOX2* c.849C>A, five of 24 sequenced clones (~21%) were confirmed to have lost the patient mutation, matching the expected frequency of 20%–26%. However, in two of them, additional mutations were introduced during the editing process. For *SHOX2* c.*28T, five of the annotated WT clones (1× from sib selection gRNA-1 no. 1 and 4× gRNA-2 no. 1) were confirmed to be isogenic with no additional detectable mutations neither ~600 bp upstream nor ~260 bp downstream of the Cas9 cut site. For the rest of the clones that were predicted to be homozygous WT via genotyping, a loss of primer/probe binding sites on the mutant allele explained the false annotation (Figures 4B and 4C).

For subsequent detailed re-characterization, we selected one isogenic clone for *SHOX2* c.849C>A and *SHOX2* c.*28T>C. Both clones had maintained their stem cell-like morphology and pluripotent capacity and exhibited high alkaline phosphatase activity, showed expression of pluripotency markers on the RNA/protein level, and spontaneously differentiated into derivates of all three germ layers (Figure S4A–D). Classical cytogenetic analysis on Giemsa-stained chromosomes revealed a normal male karyotype

(C) Screening for isogenic clones derived from heterozygous *SHOX2* c.*28T>C cells: single-cell-derived clones were genotyped. Ten annotated homozygous WT clones were sequenced to confirm the loss of the *SHOX2* c.*28T>C mutation. In 5/10 clones the mutation was repaired precisely back to WT, in the other 5/10 clones, deletions on the Mut allele explained the false annotation. Abbreviations: SNP, single-nucleotide polymorphism, here, *SHOX2* c.849C>A and *SHOX2* c.*28T>C; Indel, insertion/deletion.
See also Figures S3 and S4, Tables S3 and S5.

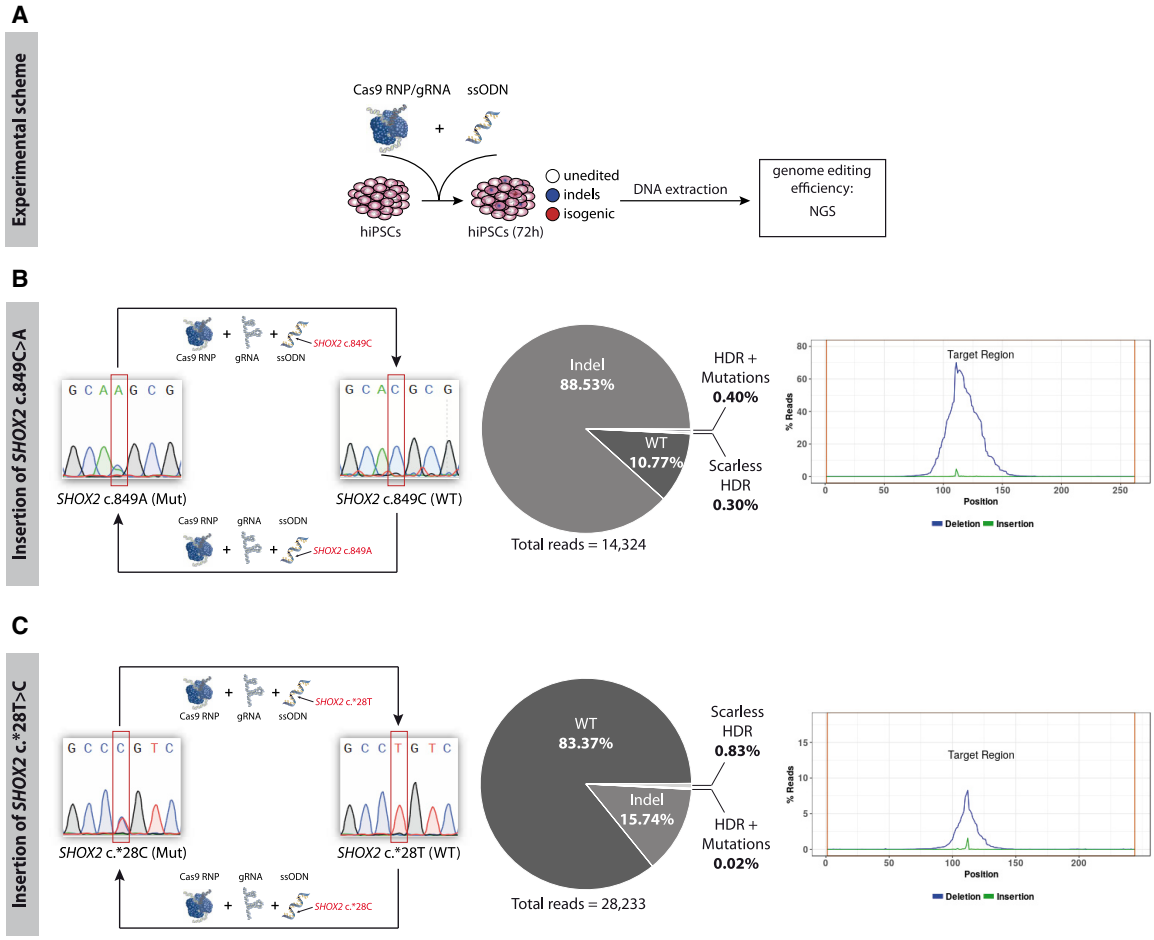


Figure 5. Reverse Genome Editing to Determine the Initial HDR Efficiency

(A) Experimental scheme for reverse gene editing. The newly generated isogenic clones were used to determine the initial HDR efficiency. For this, the same Cas9 RNP and gRNA were used in combination with an oligonucleotide donor differing only in the one base analogous to *SHOX2* c.849 (B) or *SHOX2* c.*28 (C). In consequence, HDR events would lead to the introduction of the *SHOX2* c.849C>A or *SHOX2* c.*28T>C mutation. The frequency of HDR was determined by NGS 72 h after isogenic cells were transfected with Cas9/gRNA/ssODNs. Abbreviations: HDR, homology-directed repair.

See also Tables S2, S5, and S6.

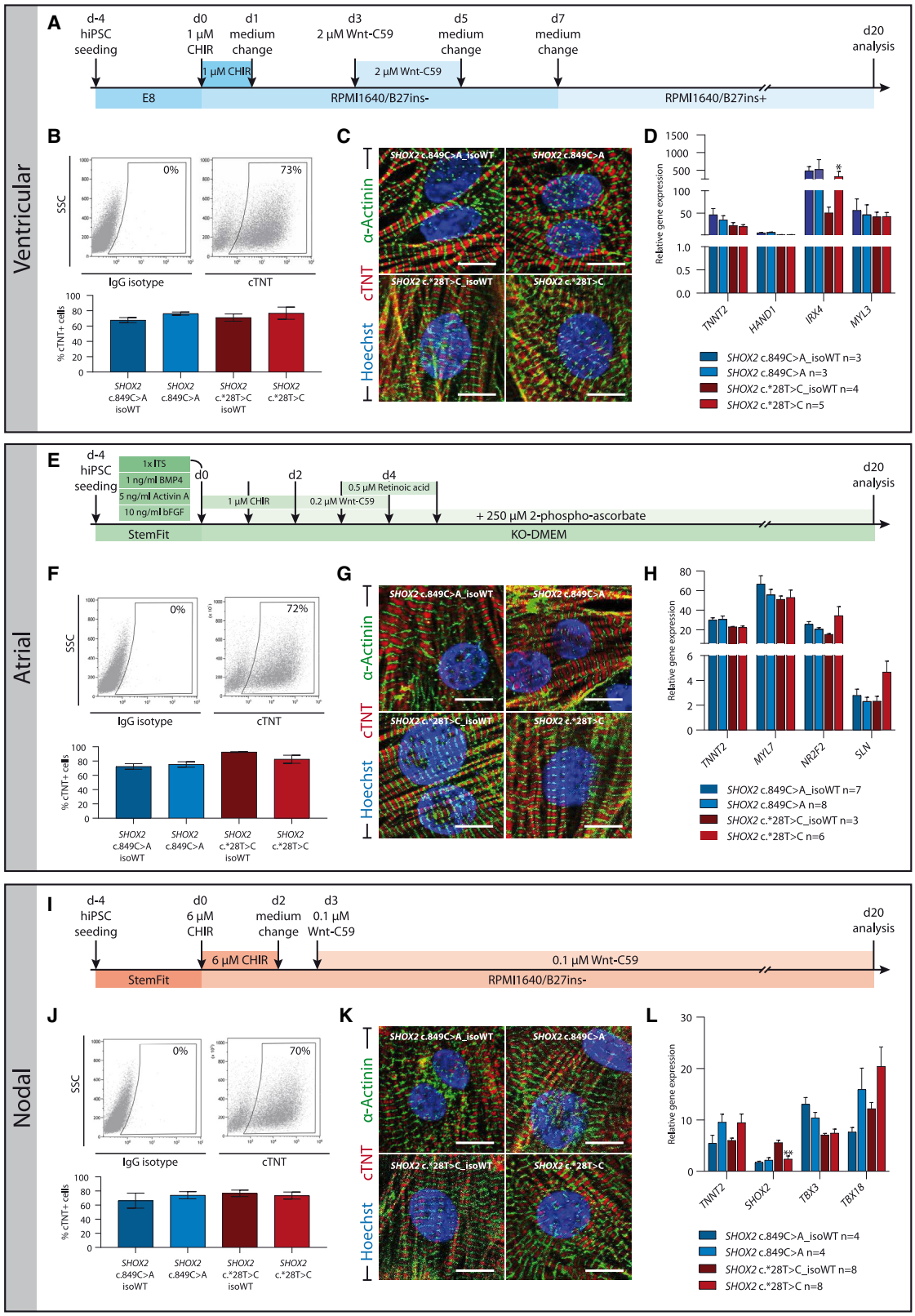
in 30 investigated metaphases each (Figure S4E). Cell line authentication confirmed their patient-specific origin and excluded cross-contamination with the control line (Table S2). We sequenced 12 and 11 highly scored off-targets for *SHOX2* c.849C>A gRNA-2 and *SHOX2* c.*28T>C gRNA-1 consisting of exonic regions as well as intronic and intergenic regions with potential regulatory relevance, but found no additional mutations introduced in these sites (Table S4).

The Initial HDR Frequency Is Determined by Re-introducing the Patient-Specific Mutations into the Isogenic Control Lines

To determine if the sib-selection process indeed enriched the targeted cells, the initial HDR efficiency had to be quantified. Therefore, *SHOX2* c.849C>A or *SHOX2* c.*28T>C mutations were re-introduced into the generated cell line using the same Cas9 enzyme batch and gRNAs but replacing the ssODNs with analogous versions that would lead to the insertion of the point mutations rather than to their repair. Seventy-two hours after transfection, the frequency of HDR events was determined by NGS. The *SHOX2* c.849C>A and *SHOX2* c.*28T>C mutations could be detected in 0.70% and 0.85% of the reads, respectively, correlating with often-reported frequencies of ~1% (Figure 5) (Miyaoka et al., 2014).

Patient-Specific and Isogenic hiPSCs Were Differentiated into Different Cardiomyocyte Subtypes

To prove the suitability of these lines as a cardiac disease model, patient-derived hiPSCs and their isogenic



(legend on next page)

counterparts were differentiated into ventricular-, atrial-, and nodal-like cardiomyocytes using previously described protocols (Foo et al., 2018; Liang et al., 2019; Zhang et al., 2015a) (Figures 6A, 6E, and 6I). Flow-cytometry analysis of the pan-cardiac marker cardiac Troponin T showed comparable differentiation capacities between isogenic pairs (Figures 6B, 6F, and 6J). Immunostaining of cardiac Troponin T (A band) and sarcomeric α-actinin (Z band) of 24-day-old cardiomyocytes revealed normal sarcomeric organization in all lines (Figures 6C, 6G, and 6K). We analyzed the expression of several subtype-specific marker genes on the RNA level and found no significant differences between patient and isogenic lines, with the exception of *SHOX2* expression that was reduced in nodal-like cells and a reduced expression of the ventricular homeobox transcription factor *IRX4* in the presence of the c.*28T>C mutation (Figures 6D, 6H, and 6L). We conclude that these lines are well suited to future investigations on how these *SHOX2* mutations contribute to the prevalence of AF.

DISCUSSION

The concept of using stochastic enrichment of cells by sib selection to introduce precise mutations into hiPSCs derived from a healthy donor has been proposed before (Miyaoka et al., 2014). The insertion of disease-linked variants into a WT background helps to interrogate the influence of these mutations on the onset and progression of a disease by direct comparison of mutated cells to their isogenic counterpart. Yet, this approach is limited to monogenic diseases or variances with a high impact on the phenotype. Here, we demonstrate that this approach can be used to correct heterozygous mutations as well. This is of particular interest when using patient-derived hiPSCs that already harbor putative disease-causing variants. These patient models play a central role in the investigation of sporadic or idiopathic diseases, where a combination of multiple risk alleles with low effect size is thought to be the genetic basis and individual risk variants might not be sufficient to cause a disease-associated phenotype.

AF is a multifactorial disease with a strong genetic component and a complex heritability (Feghaly et al., 2018). Multiple genetic loci have been associated with this disease (Roselli et al., 2018) and mutations in potassium and sodium channels as well as mutations in transcription factors and structural proteins have been identified (Feghaly et al., 2018). However, AF models based on human stem cell-derived cardiomyocytes are sparse (Benzoni et al., 2019; Laksman et al., 2017) and to date, linking specific mutations to an AF phenotype in hiPSC-cardiomyocytes has not been achieved. The correction of putative disease-contributing variants in hiPSCs could unravel subtle phenotypic changes, even if the disease phenotype overall persisted.

Our proposed strategy can be applied to a precise repair of heterozygous mutations that does not require the use of selection marker integration, its transient expression, or an enrichment of nuclease-expressing cells by fluorescence-activated cell sorting (Lonowski et al., 2017; Mitzelfelt et al., 2017; Steyer et al., 2018). This does not only abolish the need to optimize the selection or sorting process, but also allows the use of unlabeled nuclease proteins or gene constructs. The enrichment of cells with a high expression of Cas9 via puromycin or cell sorting has led to concerns regarding off-target effects, as prolonged expression or high concentrations of nucleases tend to increase the probability of unwanted DSBs (Chen et al., 2016; Zhang et al., 2015b). This can be countered by the use of Cas9 RNPs, which reportedly show less off-targeting due to the shortened activity span by immediate DNA cleavage after delivery into cells, followed by a rapid degradation (Kim et al., 2014; Ramakrishna et al., 2014). Avoiding delivery vectors that have the potential to integrate into the genome also opens possibilities for gene editing in clinical settings under good manufacturing practice conditions. In addition, high-fidelity gRNAs with few predicted off-targets should be preferentially chosen (Hsu et al., 2013). However, to increase the frequency of HDR events, gRNAs must be selected according to the distance from the induced DSB to the HDR-targeted DNA section, rather than off-target scores or predicted efficiency. Nevertheless, to completely rule out

Figure 6. Differentiation of Patient-Specific and Isogenic Control hiPSCs into Cardiomyocyte Subtypes

(A, E, and I) hiPSCs were differentiated into ventricular- (A–D), atrial- (E–H), and nodal-like (I–L) cardiomyocytes using the given differentiation schemes. (B, F, and I) Flow-cytometry analysis revealed similar numbers of $cTNT^+$ cells in isogenic pairs, indicating an unimpaired differentiation capacity in patient lines. Top: Representative flow-cytometry dot plots for IgG isotype control (left) and cTNT (right) of differentiated cells at day 20. Bottom: Percentage of $cTNT^+$ cells in n = 3 independent experiments. (C, G, and K) Immunostaining for cardiac Troponin T (sarcomere A band) and α-actinin A (sarcomere Z band) showed normal sarcomeric organization in patient and control lines; scale bar, 50 μm. (D, H, and L) qRT-PCR analysis of subtype-specific markers showed equal gene expression between patient and control hiPSC-derived cardiomyocytes, with the exception of *SHOX2* and *IRX4*, which were significantly downregulated in the *SHOX2* mutant c.*28T>C line. Data are expressed as mean ± SD of three independent experiments. *$p \leq 0.05$, **$p \leq 0.01$. Abbreviation: cTNT, cardiac Troponin T.

See also Tables S5 and S6.

additional editing, a genome-wide analysis via whole-genome sequencing would be required. Enriching cells of interest before single-cell seeding is especially beneficial for cell lines that behave poorly during clonal expansion. In our experience, keeping cells in small pools of 200 cells per well of a 96-well plate has significantly improved survival rates upon splitting. This avoids the selective enrichment of cell populations with abnormal survival advantages or growth rates caused by chromosomal aberrations. Subsequently, low clonability rates do not lead to an immense increase in time and material consumption due to large-scale cloning efforts as only a handful of clones have to be analyzed to find an isogenic one. In fact, even with the low clonability rates of 10%–20% that we achieved with our patient lines, only two or three 96-well plates per line were sufficient to find several isogenic populations.

In regions that are difficult to target, the initial HDR efficiency may still be low even after incorporation of improvement strategies such as optimizing the DNA template (Okamoto et al., 2019; Richardson et al., 2016), modifying the Cas9 enzyme itself (Anzalone et al., 2019; Komor et al., 2016), or applying small molecules to increase HDR (Yu et al., 2015). Our proposed strategy can be combined with any of these approaches as it is locus independent and relies solely on a stochastic enrichment rather than the modulation of biological processes. Especially, emerging techniques like base editing and prime editing (Anzalone et al., 2019; Komor et al., 2016) can profit from this approach by complementing their increased efficiency in targeted editing with an additional enrichment of correctly edited cells. We purposely avoided the insertion of silent blocking mutations, which prevent re-cutting by Cas9 after HDR-mediated editing (Okamoto et al., 2019), together with the mutation correction in the ssODN sequence. We highly recommend restricting their application to coding mutations, as unwanted side effects in UTRs or non-coding regulatory regions by altered posttranscriptional regulation or transcription factor binding cannot be ruled out completely. An example of this is the *SHOX2* c.*28T>C mutation, which resides within the 3′ UTR and mediates its detrimental effect by the generation of a novel microRNA binding site (Hoffmann et al., 2016).

However, despite the clear advantages of this method, several limitations remain. Laboratories are required to have fast and easy access to NGS and potentially dPCR. We used dPCR to preselect sib selections with potentially high fractions of isogenic cells, despite the stated limitations of this method to quantify such subpopulations. In addition, some dPCR primer/probe combinations seem to systematically favor the amplification of WT or Mut alleles in dPCRs, thus allowing only a relative comparison between sib selections. The preselection of a few samples made singleplex NGS possible, which could be analyzed with freely available online tools such as Cas-analyzer (Park et al., 2017). Nevertheless, a more straightforward approach would be to quantify alleles in all sib selections at once with deep sequencing. Although this is more cost intense and requires amplicon-NGS analysis knowledge due to sample multiplexing, it would speed up the process significantly and allow one to find the one sib selection that truly has the highest percentage of isogenic cells. With recent and future advances in NGS, we expect this method to become more feasible and affordable (Levy and Myers, 2016; Park and Kim, 2016). The additional passaging and cryopreservation required in the sib-selection process is a potential source for acquiring mutations and chromosomal translocations in the extended culturing periods (Martins-Taylor and Xu, 2012; Merkle and Eggan, 2013). On the other hand, stressful processes like cell sorting and antibiotic selection are not required. Yet, its applicability to primary cells and cell types that cannot be extensively passaged is limited. Furthermore, the additional culturing periods are time consuming. Even under ideal circumstances, the isolation of isogenic clones takes several weeks to months. However, most of the time is spent waiting for the cells to grow, and during allele quantification procedures, sib selections are cryopreserved, making the process easily interruptible. The quantification of HDR events after each sib selection allows precise workload calculations for the next step and prevents tedious single-cell cloning, even if no isogenic cells are present.

In conclusion, we provide a strategy for the scarless correction of heterozygous mutations by random enrichment of precisely edited cells and their detection via allele quantification. We propose that the frequency of isogenic cells can be determined by comparing WT/Mut allele ratios with two assumptions: no CNV and an error-free correction of the mutation via HDR. This approach can facilitate the generation of isogenic control cells, which represent the gold standard of controls when investigating the influence of putative disease-causing variants on the disease phenotype. Future detailed electrophysiological and molecular analysis of this *in vitro* disease model will aid in refining the paradigm of how *SHOX2* mutations influence the onset and progression of AF by direct genotype-phenotype comparison and even has the potential to serve as a platform for drug discovery and personalized medicine. Compared with already existing AF models (Benzoni et al., 2019; Marczenke et al., 2017), this would be one using patient-specific and gene-corrected hiPSCs.

EXPERIMENTAL PROCEDURES

Cell Culture and Reagents

Standard cultivation of hiPSCs was performed with 5% CO_2 at 37°C in a humidified atmosphere. Cells were grown in StemFit, supplemented with 100 ng/mL basic fibroblast growth factor

(Peprotech) on a Geltrex LDEV-free, hESC-qualified, Reduced Growth Factor Basement Membrane matrix (1:100, Thermo Fisher) according to the manufacturer's instructions. Cells were dissociated with TrypLE Express (Gibco).

CRISPR/Cas9 gRNA Design, Off-Target Analysis, and gRNA Synthesis

gRNA design was performed with the *CCTop-CRISPR/Cas9 target online predictor* (https://crispr.cos.uni-heidelberg.de/) (Stemmer et al., 2015) with spCas9, a target-site length of 20 nucleotides, and standard settings as input (Table S5). Off-targets were predicted with *CCTop* using standard settings and the *Homo sapiens* GRCh37/hg19 assembly as a reference genome. The top 20 off-target sites 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, 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 (12 and 11 off-targets per gRNA) (Table S3). gRNAs were *in vitro* synthesized with the *GeneArt Precision gRNA Synthesis Kit* (Invitrogen) according to the provided protocol and quantified with the Qubit RNA BR Assay Kit (Invitrogen) on the Qubit 2.0 fluorometer (Invitrogen).

Design of Primers/Probes and Oligonucleotide Donors

ssODNs were designed as previously described (Bollen et al., 2018; Richardson et al., 2016). The template spanned the Cas9 cut site asymmetrically with a longer 5′ homology arm for improved hybridization (Table S5). The DNA oligo was synthesized and purified via desaltation by Integrated DNA Technologies. Primers and probes for the dPCR were designed via Primer3plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) and further evaluated with OligoAnalyzer (http://eu.idtdna.com/calc/analyzer) and *UCSC in silico PCR* (http://rohsdb.cmb.usc.edu/GBshape/cgi-bin/hgPcr). Primers were designed by default for an annealing temperature of 60°C and TaqMan probes for 65°C. The PCR product length was set to be 90–120 nucleotides for dPCR or 150–450 nucleotides for NGS (Table S5). For quantitative PCR, primers were designed with the *Universal Probe-Library Assay Design Center* (Roche) (https://lifescience.roche.com/en_de/brands/universal-probe-library.html#assay-design-center) (Table S5).

Electroporation and Sib Selection

iPSCs were dissociated with TrypLE Express and transfected with the 4D-Nucleofector System (Lonza) and the *P3 Primary Cell 4D-Nucleofector X Kit S* according to manufacturer's guidelines. For genome editing, 1 μg of Platinum or Truecut v.2 Cas9 RNP (Invitrogen), 250 ng of *in vitro* synthesized gRNA, and 50 pmol of ssODN were electroporated into 200,000 hiPSCs using program DN-100. Cells were seeded at high density posttransfection (50,000 cells/cm^2) and cultured in medium containing 10 μM Y-27632 (ROCK inhibitor) for the first 24 h Forty-eight hours after transfection, hiPSCs were dissociated and seeded at 200 cells/well on a precoated 96-well plate (sib selection). The amount of 200 cells per sib selection was chosen as a compromise between cell viability upon splitting and throughput (the more cells per well, the better) and a cell pool size that is small enough to also detect smaller isogenic subpopulations. In a sib selection of 200 cells, the minimum allele frequency is 0.25%. If an overaverage abundance of isogenic cells (>1% or >2/200 cells) is present, the shift in the allele frequency would be >0.5%, which we considered to be detectable in dPCR and NGS. These sib selections were cultured until they reached ~80% confluency (8–10 days) and then dissociated with 30 μL TrypLE Express per well. Fifteen microliters of the cell suspension was directly used for DNA extraction with the *Quick*-DNA 96 Plus Kit (Zymo Research), the other 15 μL was mixed with 85 μL Bambanker (Nippon Genetics) and frozen in a Styrofoam container at −80°C until further use.

dPCR

The dPCR consisted of 250 nM allele-specific TaqMan probes labeled with FAM (Mut allele) or HEX (WT allele) and 900 nM forward + reverse primers in 1× dPCR Supermix for Probes (Bio-Rad), containing 50–150 ng of genomic DNA or 50,000 plasmid copies in a total volume of 22 μL before the generation of droplets was performed with a QX200 droplet generator according to the manufacturer's instructions (Bio-Rad). Droplets were immediately pipetted into 96-well PCR plates (Bio-Rad), which were heat sealed and transferred into a C1000 thermal cycler (Bio-Rad). The conducted thermal cycling program consisted of step 1, 95°C 10 min; step 2, 94°C 30 s; step 3, 57°C or 60°C 60 s, repetition of step 2 + 3 39 times, 98°C 10 min. After completion, the droplets were analyzed on a QX200 droplet reader (Bio-Rad) using the Quantasoft software (version 1.7.4.0917). To calculate the percentage of WT alleles, WT allele copies/mL was divided by WT plus Mut copies/mL. We chose sib selections for NGS based on the percentage of WT alleles and the number of detectable copies/μL (value given by the droplet reader). We hypothesized that a high value here indicates a reliable result and a dPCR that was not inhibited by non-detectable DNA alleles, thus giving a small range of possible allele ratios.

Next-Generation Sequencing and Analysis

For amplicon-NGS, the target regions spanning the *SHOX2* mutations and the Cas9 cut site were amplified using the Q5 High-Fidelity DNA polymerase (New England Biolabs) or the HotStarTaq polymerase (QIAGEN) according to the manufacturer's instructions. Five hundred nanograms of the PCR product in 25 μL was sent to GENEWIZ, Germany, for single-sample paired-end amplicon NGS on the HiSeq system (Illumina). To determine gRNA efficiencies and the frequency of HDR upon reverse gene editing, the provided SNP/indel analysis from GENEWIZ was used. For allele quantification in sib selections, the .fastq files were uploaded to the JavaScript-based assessment tool Cas-Analyzer (http://www.rgenome.net/cas-analyzer/) with the following analysis parameters: comparison range 100 nucleotides, minimum frequency 0.25% of total reads containing both indicator sequences (determined by setting it to 1 at first, doing the analysis, and then calculating 0.25% from the total number of reads containing both

indicator sequences) and no WT marker. The resulting sequence table was used for manual allele classification (Table S4).

Single-Cell Cloning

Thawed sib selections were dissociated and a cell suspension with 10 cells/mL was generated by serial dilution. The suspension was mixed 1:50 with ice-cold Geltrex and 100 μL per well (=1 cell/well) was immediately pipetted on a 96-well plate using multichannel pipets. Seventy-two hours after seeding, plates were screened for single cell-derived clones (Figure S3). Medium was changed every other day until day 8, when hiPSC colonies were split onto 2× precoated 96 wells each. After reaching confluency, one well was subjected to cell lysis with directPCR lysis reagent (VWR), the other was used for expansion and cryopreservation.

Genotyping

To identify potential isogenic WT clones, single-cell-derived clones were genotyped in 96-well plates via TaqMan probe-based SNP genotyping using 1× TaqPath ProAmp Master Mix (Thermo Fisher) according to the manufacturer's instructions. DNA from control and patient-specific cells served as positive control for homozygous and heterozygous WT annotations. Potential isogenic clones were identified in the allelic discrimination plot.

Subtype-Specific Cardiomyocyte Differentiation

Differentiation into ventricular- (Foo et al., 2018), atrial- (Zhang et al., 2015a), and nodal-like cardiomyocytes (Liang et al., 2019) was done as described previously. At day 20, differentiating myocytes were dissociated with papain (Worthington Biochemical Corporation) and analyzed for $cTNT^+$ cells by flow cytometry or further purified with the human PSC-derived Cardiomyocyte Isolation Kit (Miltenyi Biotech) followed by RNA isolation for qPCR analyses or plating on fibronectin-coated dishes for immunohistological analysis (detailed methods are described in the Supplemental Information).

Data and Code Availability

Unprocessed sequencing files (.fastq) have been uploaded to the NCBI Sequence Read Archive and are available via the accession no. PRJNA659248 or under https://www.ncbi.nlm.nih.gov/bioproject/PRJNA659248/.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.stemcr.2020.08.015.

AUTHOR CONTRIBUTIONS

Conceptualization, S.A.S. and S.H.; Methodology, S.A.S. and S.H.; Validation, S.A.S. and S.H.; Formal Analysis, S.A.S. and S.H.; Investigation, S.A.S., S.H., K.R., S.L., B.C., K.R., V.F., T.D., J.W.G.J., and A.J.; Resources, S.C., S.K., J.W.G.J., A.J., A.M., and G.A.R.; Writing – Original Draft, S.A.S.; Writing – Review & Editing, S.H., T.D., A.M., and G.A.R.; Supervision, A.M. and G.A.R.; Funding Acquisition, S.H., A.M., K-L.L., and G.A.R.

ACKNOWLEDGMENTS

The authors declare no conflict of interest. This work was funded by the German Research Foundation (DFG) RA 380/14-4 (to G.A.R.); DZHK (German Centre for Cardiovascular Research) 81X2500133 and 81X2500113 (to S.H.); the European Research Council, ERC 788381 (to A.M.); the German Research Foundation (DFG), Transregio Research Unit 152 (to A.M., K-L.L.) and 267 (to A.M., K-L.L.); and DZHK (German Centre for Cardiovascular Research) 81X2600603 and 81X2600607 (to A.M.).

Received: March 20, 2020
Revised: August 27, 2020
Accepted: August 28, 2020
Published: September 24, 2020

REFERENCES

Abou-Saleh, H., Zouein, F.A., El-Yazbi, A., Sanoudou, D., Raynaud, C., Rao, C., Pintus, G., Dehaini, H., and Eid, A.H. (2018). The march of pluripotent stem cells in cardiovascular regenerative medicine. Stem Cell Res. Ther. *9*, 201.

Anzalone, A.V., Randolph, P.B., Davis, J.R., Sousa, A.A., Koblan, L.W., Levy, J.M., Chen, P.J., Wilson, C., Newby, G.A., Raguram, A., and Liu, D.R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. Nature *576*, 149–157.

Benzoni, P., Campostrini, G., Landi, S., Bertini, V., Marchina, E., Iascone, M., Ahlberg, G., Olesen, M.S., Crescini, E., Mora, C., et al. (2019). Human iPSC modeling of a familial form of atrial fibrillation reveals a gain of function of if and ICaL in patient-derived cardiomyocytes. Cardiovasc. Res. *116*, 1147–1160.

Blaschke, R.J., Hahurij, N.D., Kuijper, S., Just, S., Wisse, L.J., Deissler, K., Maxelon, T., Anastassiadis, K., Spitzer, J., Hardt, S.E., et al. (2007). Targeted mutation reveals essential functions of the homeodomain transcription factor Shox2 in sinoatrial and pacemaking development. Circulation *115*, 1830–1838.

Bollen, Y., Post, J., Koo, B.K., and Snippert, H.J.G. (2018). How to create state-of-the-art genetic model systems: strategies for optimal CRISPR-mediated genome editing. Nucleic Acids Res. *46*, 6435–6454.

Brandao, K.O., Tabel, V.A., Atsma, D.E., Mummery, C.L., and Davis, R.P. (2017). Human pluripotent stem cell models of cardiac disease: from mechanisms to therapies. Dis. Model. Mech. *10*, 1039–1059.

Brookhouser, N., Raman, S., Potts, C., and Brafman, D.A. (2017). May i cut in? gene editing approaches in human induced pluripotent stem cells. Cells *6*, 5.

Chen, Y., Liu, X., Zhang, Y., Wang, H., Ying, H., Liu, M., Li, D., Lui, K.O., and Ding, Q. (2016). A self-restricted CRISPR system to reduce off-target effects. Mol. Ther. *24*, 1508–1510.

Chu, V.T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., and Kuhn, R. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat. Biotechnol. *33*, 543–548.

El-Battrawy, I., Zhao, Z., Lan, H., Li, X., Yucel, G., Lang, S., Sattler, K., Schunemann, J.D., Zimmermann, W.H., Cyganek, L., et al. (2018). Ion channel dysfunctions in dilated cardiomyopathy in

limb-girdle muscular dystrophy. Circ. Genom Precis Med. *11*, e001893.

Elliott, B., Richardson, C., Winderbaum, J., Nickoloff, J.A., and Jasin, M. (1998). Gene conversion tracts from double-strand break repair in mammalian cells. Mol. Cell. Biol. *18*, 93–101.

Feghaly, J., Zakka, P., London, B., Macrae, C.A., and Refaat, M.M. (2018). Genetics of atrial fibrillation. J. Am. Heart Assoc. *7*, e009884.

Foo, K.S., Lehtinen, M.L., Leung, C.Y., Lian, X., Xu, J., Keung, W., Geng, L., Kolstad, T.R.S., Thams, S., Wong, A.O., et al. (2018). Human ISL1(+) ventricular progenitors self-assemble into an in vivo functional heart patch and preserve cardiac function post infarction. Mol. Ther. *26*, 1644–1659.

Goedel, A., MY, I., Sinnecker, D., and Moretti, A. (2017). Perspectives and challenges of pluripotent stem cells in cardiac arrhythmia research. Curr. Cardiol. Rep. *19*, 23.

Guo, Q., Mintier, G., Ma-Edmonds, M., Storton, D., Wang, X., Xiao, X., Kienzle, B., Zhao, D., and Feder, J.N. (2018). 'Cold shock' increases the frequency of homology directed repair gene editing in induced pluripotent stem cells. Sci. Rep. *8*, 2080.

Hoffmann, S., Berger, I.M., Glaser, A., Bacon, C., Li, L., Gretz, N., Steinbeisser, H., Rottbauer, W., Just, S., and Rappold, G. (2013). Islet1 is a direct transcriptional target of the homeodomain transcription factor Shox2 and rescues the Shox2-mediated bradycardia. Basic Res. Cardiol. *108*, 339.

Hoffmann, S., Clauss, S., Berger, I.M., Weiss, B., Montalbano, A., Roth, R., Bucher, M., Klier, I., Wakili, R., Seitz, H., et al. (2016). Coding and non-coding variants in the SHOX2 gene in patients with early-onset atrial fibrillation. Basic Res. Cardiol. *111*, 36.

Hoffmann, S., Paone, C., Sumer, S.A., Diebold, S., Weiss, B., Roeth, R., Clauss, S., Klier, I., Kaab, S., Schulz, A., et al. (2019). Functional characterization of rare variants in the SHOX2 gene identified in sinus node dysfunction and atrial fibrillation. Front. Genet. *10*, 648.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. *31*, 827–832.

Jackson, S.P., and Bartek, J. (2009). The DNA-damage response in human biology and disease. Nature *461*, 1071–1078.

Jang, Y.Y., and Ye, Z. (2016). Gene correction in patient-specific iPSCs for therapy development and disease modeling. Hum. Genet. *135*, 1041–1058.

Kim, S., Kim, D., Cho, S.W., Kim, J., and Kim, J.S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res. *24*, 1012–1019.

Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature *533*, 420–424.

Laksman, Z., Wauchop, M., Lin, E., Protze, S., Lee, J., Yang, W., Izaddoustdar, F., Shafaattalab, S., Gepstein, L., Tibbits, G.F., et al. (2017). Modeling atrial fibrillation using human embryonic stem cell-derived atrial tissue. Sci. Rep. *7*, 5268.

Levy, S.E., and Myers, R.M. (2016). Advancements in next-generation sequencing. Annu. Rev. Genomics Hum. Genet. *17*, 95–115.

Li, N., Wang, Z.S., Wang, X.H., Xu, Y.J., Qiao, Q., Li, X.M., Di, R.M., Guo, X.J., Li, R.G., Zhang, M., et al. (2018). A SHOX2 loss-of-function mutation underlying familial atrial fibrillation. Int. J. Med. Sci. *15*, 1564–1572.

Liang, W., Han, P., Kim, E.H., Mak, J., Zhang, R., Torrente, A.G., Goldhaber, J.I., Marban, E., and Cho, H.C. (2019). Canonical Wnt signaling promotes pacemaker cell specification of cardiac mesodermal cells derived from mouse and human embryonic stem cells. Stem Cells *38*, 352–368.

Lonowski, L.A., Narimatsu, Y., Riaz, A., Delay, C.E., Yang, Z., Niola, F., Duda, K., Ober, E.A., Clausen, H., Wandall, H.H., et al. (2017). Genome editing using FACS enrichment of nuclease-expressing cells and indel detection by amplicon analysis. Nat. Protoc. *12*, 581–603.

Marczenke, M., Fell, J., Piccini, I., Ropke, A., Seebohm, G., and Greber, B. (2017). Generation and cardiac subtype-specific differentiation of PITX2-deficient human iPS cell lines for exploring familial atrial fibrillation. Stem Cell Res. *21*, 26–28.

Martins-Taylor, K., and Xu, R.H. (2012). Concise review: genomic stability of human induced pluripotent stem cells. Stem Cells *30*, 22–27.

Merkle, F.T., and Eggan, K. (2013). Modeling human disease with pluripotent stem cells: from genome association to function. Cell Stem Cell *12*, 656–668.

Mitzelfelt, K.A., Mcdermott-Roe, C., Grzybowski, M.N., Marquez, M., Kuo, C.T., Riedel, M., Lai, S., Choi, M.J., Kolander, K.D., Helbling, D., et al. (2017). Efficient precision genome editing in iPSCs via genetic co-targeting with selection. Stem Cell Rep. *8*, 491–499.

Miyaoka, Y., Chan, A.H., Judge, L.M., Yoo, J., Huang, M., Nguyen, T.D., Lizarraga, P.P., So, P.L., and Conklin, B.R. (2014). Isolation of single-base genome-edited human iPS cells without antibiotic selection. Nat. Methods *11*, 291–293.

Moretti, A., Fonteyne, L., Giesert, F., Hoppmann, P., Meier, A.B., Bozoglu, T., Baehr, A., Schneider, C.M., Sinnecker, D., Klett, K., et al. (2020). Somatic gene editing ameliorates skeletal and cardiac muscle failure in pig and human models of Duchenne muscular dystrophy. Nat. Med. *26*, 207–214.

Moretti, A., Laugwitz, K.L., Dorn, T., Sinnecker, D., and Mummery, C. (2013). Pluripotent stem cell models of human heart disease. Cold Spring Harb. Perspect. Med. *3*, a014027.

Okamoto, S., Amaishi, Y., Maki, I., Enoki, T., and Mineno, J. (2019). Highly efficient genome editing for single-base substitutions using optimized ssODNs with Cas9-RNPs. Sci. Rep. *9*, 4811.

Park, J., Lim, K., Kim, J.S., and Bae, S. (2017). Cas-analyzer: an online tool for assessing genome editing results using NGS data. Bioinformatics *33*, 286–288.

Park, S.T., and Kim, J. (2016). Trends in next-generation sequencing and a new era for whole genome sequencing. Int. Neurourol J. *20*, S76–S83.

Ramakrishna, S., Kwaku Dad, A.B., Beloor, J., Gopalappa, R., Lee, S.K., and Kim, H. (2014). Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. Genome Res. *24*, 1020–1027.

Richardson, C.D., Ray, G.J., Dewitt, M.A., Curie, G.L., and Corn, J.E. (2016). Enhancing homology-directed genome editing by

catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat. Biotechnol. *34*, 339–344.

Roselli, C., Chaffin, M.D., Weng, L.C., Aeschbacher, S., Ahlberg, G., Albert, C.M., Almgren, P., Alonso, A., Anderson, C.D., Aragam, K.G., et al. (2018). Multi-ethnic genome-wide association study for atrial fibrillation. Nat. Genet. *50*, 1225–1233.

Soldner, F., Laganiere, J., Cheng, A.W., Hockemeyer, D., Gao, Q., Alagappan, R., Khurana, V., Golbe, L.I., Myers, R.H., Lindquist, S., et al. (2011). Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. Cell *146*, 318–331.

Stemmer, M., Thumberger, T., Del Sol Keyer, M., Wittbrodt, J., and Mateo, J.L. (2015). CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. PLoS One *10*, e0124633.

Steyer, B., Bu, Q., Cory, E., Jiang, K., Duong, S., Sinha, D., Steltzer, S., Gamm, D., Chang, Q., and Saha, K. (2018). Scarless genome editing of human pluripotent stem cells via transient puromycin selection. Stem Cell Rep. *10*, 642–654.

van Mil, A., Balk, G.M., Neef, K., Buikema, J.W., Asselbergs, F.W., Wu, S.M., Doevendans, P.A., and Sluijter, J.P.G. (2018). Modelling inherited cardiac disease using human induced pluripotent stem cell-derived cardiomyocytes: progress, pitfalls, and potential. Cardiovasc. Res. *114*, 1828–1842.

Ye, W., Wang, J., Song, Y., Yu, D., Sun, C., Liu, C., Chen, F., Zhang, Y., Wang, F., Harvey, R.P., et al. (2015). A common Shox2-Nkx2-5 antagonistic mechanism primes the pacemaker cell fate in the pulmonary vein myocardium and sinoatrial node. Development *142*, 2521–2532.

Yu, C., Liu, Y., Ma, T., Liu, K., Xu, S., Zhang, Y., Liu, H., La Russa, M., Xie, M., Ding, S., and Qi, L.S. (2015). Small molecules enhance CRISPR genome editing in pluripotent stem cells. Cell Stem Cell *16*, 142–147.

Zhang, M., Schulte, J.S., Heinick, A., Piccini, I., Rao, J., Quaranta, R., Zeuschner, D., Malan, D., Kim, K.P., Ropke, A., et al. (2015a). Universal cardiac induction of human pluripotent stem cells in two and three-dimensional formats: implications for in vitro maturation. Stem Cells *33*, 1456–1469.

Zhang, X.H., Tee, L.Y., Wang, X.G., Huang, Q.S., and Yang, S.H. (2015b). Off-target effects in CRISPR/Cas9-mediated genome engineering. Mol. Ther. Nucleic Acids *4*, e264.

Stem Cell Reports, Volume *15*

Supplemental Information

Precise Correction of Heterozygous *SHOX2* Mutations in hiPSCs Derived 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

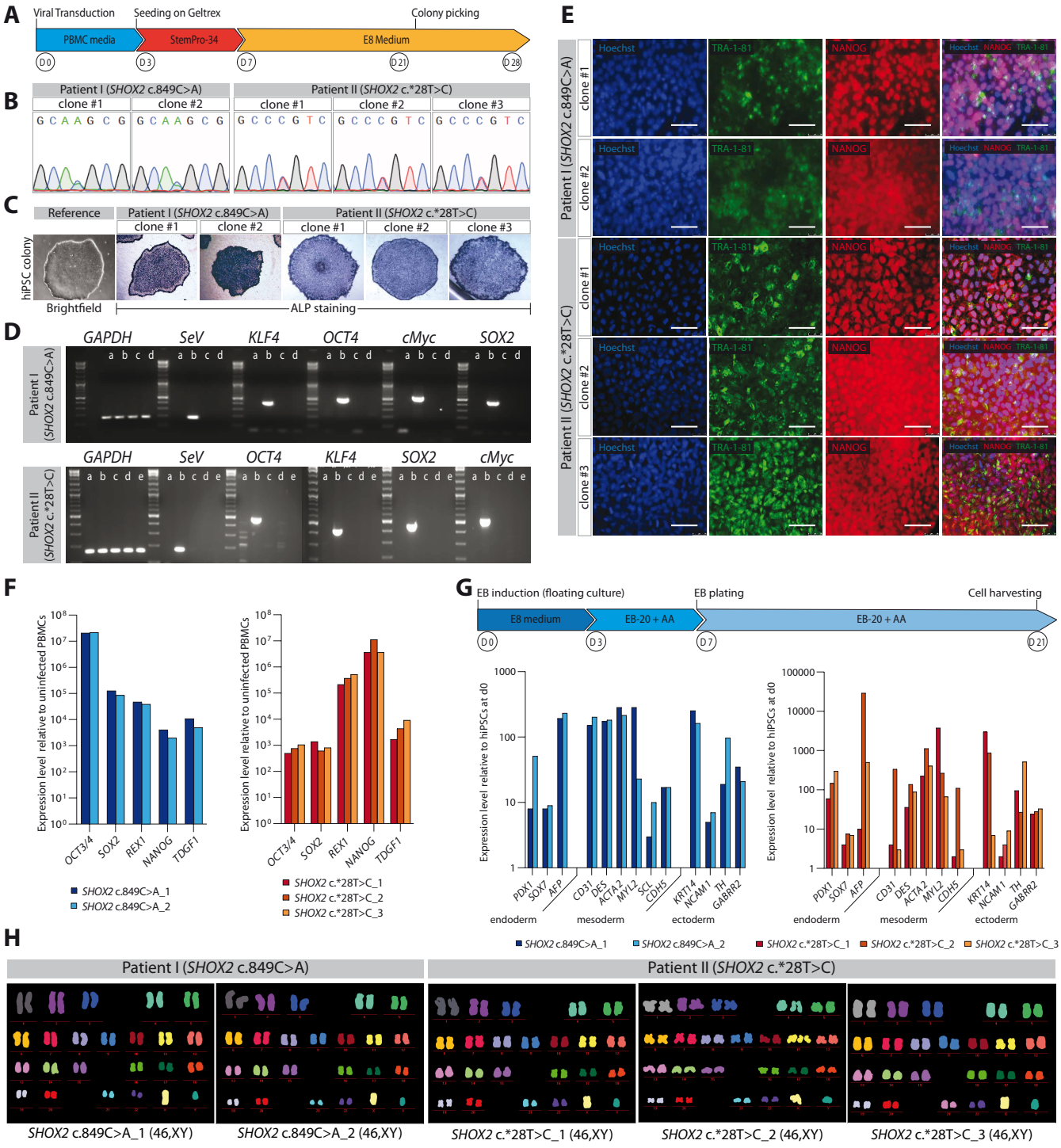


Figure S1 Generation and characterization of patient-specific hiPSCs from patient I (*SHOX2* c.849C>A) and patient II (*SHOX2* c.*28T>C). Related to Figure 1 hiPSC were generated by reprogramming peripheral blood mononuclear cells (PBMCs) using the CytoTune-hiPSC 2.0 Sendai Reprogramming Kit (Life Technologies). (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.

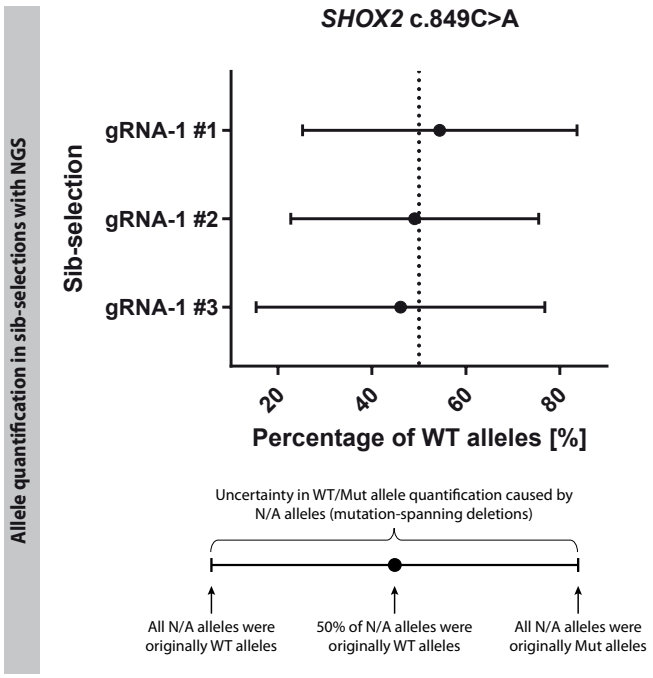


Figure S2 Allele quantification of *SHOX2* c.849C>A gRNA-1 transfected sib-selections via NGS. Related to Figure 3 Alleles with *SHOX2* c.849-spanning deletions cause an uncertainty in allele quantification that is addressed by defining those alleles as all WT or all Mut. The resulting span of possible 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

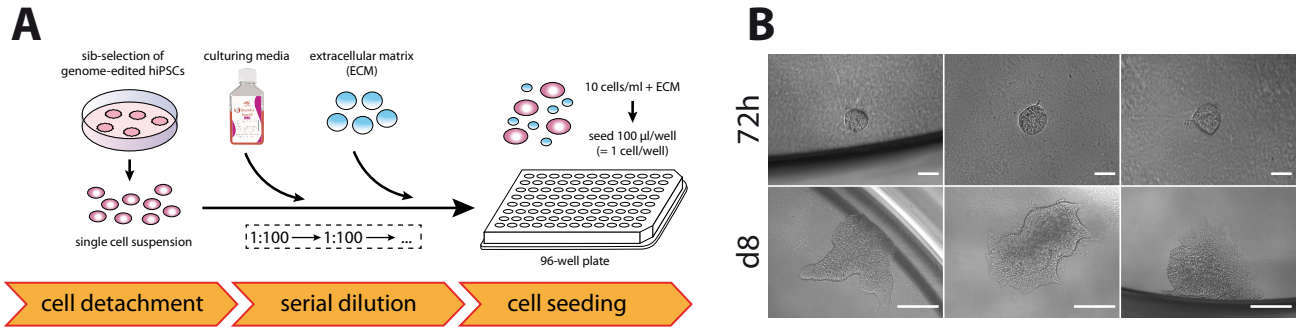


Figure S3 Single cell-cloning of hiPSCs. Related to Figure 4 (A) Coating free method for single-cell cloning via limited dilution. A single-cell suspension is generated from the chosen sib-selection and diluted to 10 cells/ml. The extracellular matrix is added to the cell suspension and 1 cell/well is 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).

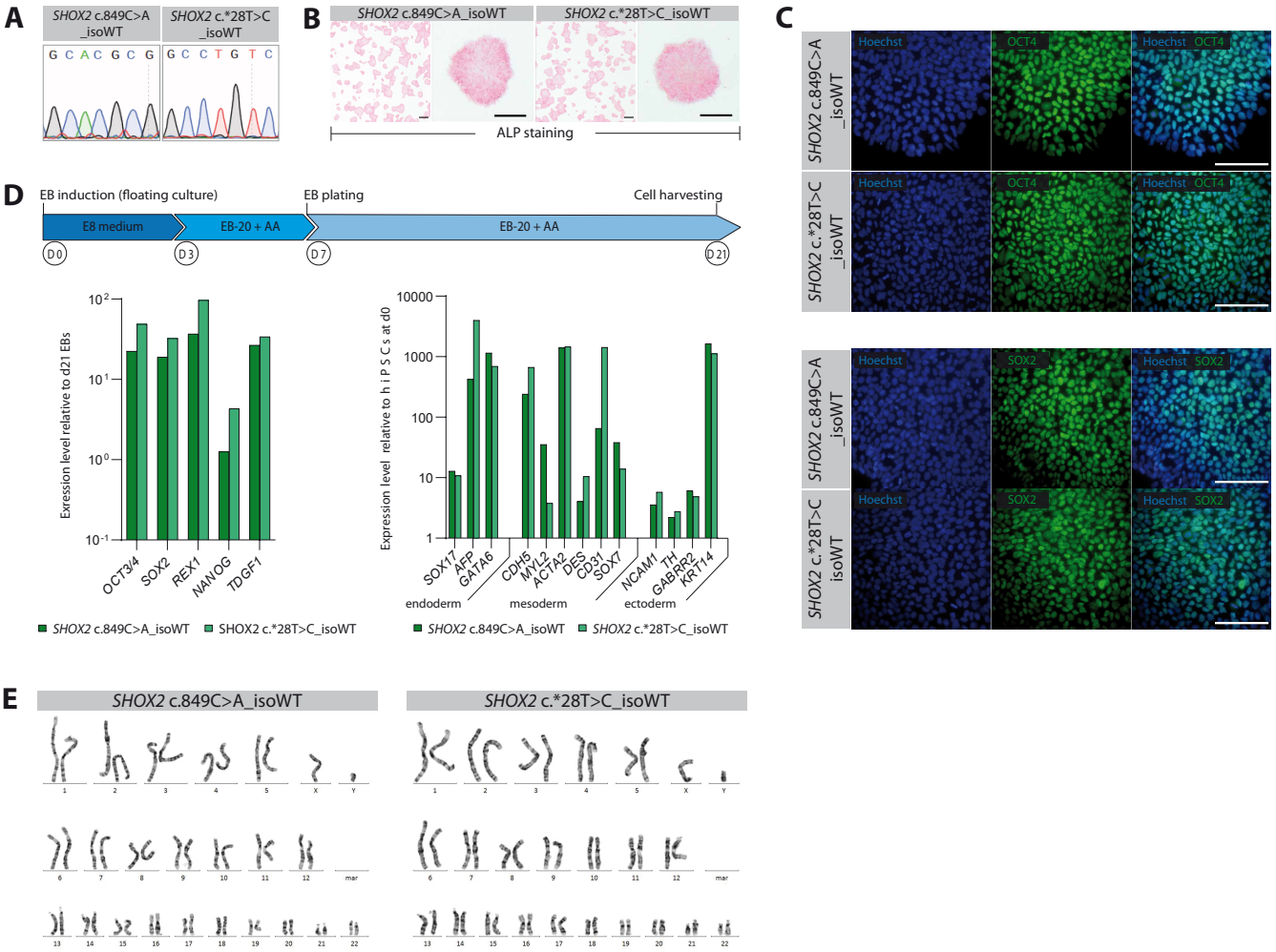


Figure S4 Re-characterization of *SHOX2* c.849C>A_isoWT and *SHOX2* c.*28T>C_isoWT. Related to Figure 4 Isogenic control lines were generated as described above and re-characterized to confirm the preservation of pluripotency capacity and karyotype.
(A) Loss of heterozygous *SHOX2* c.849C>A (= SHOX2 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 (*SHOX2* c.849C>A)	AF patient II (*SHOX2* 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
Echocardiography (at time of enrollment)	**heart transplantation**	no	yes
	ejection fraction	43%	25%
	structural	LA diameter = 43 mm	LVEED = 67 mm
Electro-cardiogram	**frequency [beats/min]**	59	71
	P wave [ms]	80	110
	PR interval [ms]	180	210
	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 TAAAAAAATATATTTTTAAAAAA
SHOX2_c.849C>A_1	hiPSCs	ok	WILL-1	90	no	no	unique sequence	ATTTATATAAATAAAATTAAAATTT TAAAAAAATAAAAAATTATWTAT
SHOX2_c.849C>A_2	hiPSCs	ok	WILL-2	90	no	no	unique sequence	ATTTATATAAATAAAATTAAAATTT TAAAAAAATAAAAAATTATWTAT
SHOX2_c.*28T>C_1	hiPSCs	ok	BSM	88	no	no	unique sequence	TTTTATATATAAATATTTTTAATTTT ATATATATAAAATTAATTTTAA
SHOX2_c.*28T>C_2	hiPSCs	ok	BSM	88	no	no	unique sequence	TTTTATATATAAATATTTTTAATTTT ATATATATAAAATTAATTTTAA
SHOX2_c.*28T>C_3	hiPSCs	ok	BSM	88	no	no	unique sequence	TTTTATATATAAATATTTTTAATTTT ATATATATAAAATTAATTTTAA
SHOX2_c.849C>A_isoWT	hiPSCs	ok	NK-92	92	no	no	unique sequence	ATTTATATAAATAAAATTAAAATTT TAAAAAAATAAAAAWTTATTTAT
SHOX2_c.*28T>C_isoWT	hiPSCs	ok	BSM	88	no	no	unique sequence	TTTTATATATAAATATTTTTAATTTT ATATATATAWAATTAATTTTAA

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.849 or *SHOX2* c.28* were deleted).

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

	Coordinates	Strand	MM	target_sequence	Distance		Gene name	Gene ID	Seq?
1	chr10:78752841-78752863	-	4	GGGACAGC[TGGCCAGCGCTG]	2179	I	KCNMA1	ENSG00000156113	ok
2	chr19:456297-456319	+	4	GCTCCTGG[GGGACAGCGCTG]	3052	I	RNA5SP462	ENSG00000252539	ok
3	chr21:46720679-46720701	+	4	GCAGGCGC[TGCACAGCGCTG]	225	I	LINC00315	ENSG00000184274	ok
4	chr8:110657693-110657715	-	4	GCAGCTGC[AGGGCAGCGCTG]	0	E	RP11-422N16.3	ENSG00000248050	ok
5	chr2:26256570-26256592	-	4	TCTCCCGC[TGGACTGCGCTG]	384	-	RAB10	ENSG00000084733	ok
6	chr20:31573388-31573410	-	4	GCTGGAAG[TGGCCAGCGCTG]	132	I	SUN5	ENSG00000167098	ok
7	chr2:242667205-242667227	+	4	GCTACCCC[TGGAAAGCGCTG]	0	E	ING5	ENSG00000168395	ok
8	chr1:151777738-151777760	+	4	ACTGCATC[TGTCCAGCGCTG]	0	E	LINGO4	ENSG00000213171	ok
9	chr20:24597308-24597330	+	3	GAGGCAGC[TGGACAGGGCTG]	19519	I	SYNDIG1	ENSG00000101463	ok
10	chr9:137772667-137772689	+	3	GATGGAGC[TGGACAGAGCTG]	0	E	FCN2	ENSG00000160339	ok
11	chr14:92673870-92673892	+	4	CCGGCAGC[TGGGGAGCGCTG]	43115	-	CPSF2	ENSG00000165934	ok
12	chr2:86260801-86260823	-	4	GCGGGATC[TGGACAACGCTG]	0	E	POLR1A	ENSG00000068654	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	I	TSHZ3	ENSG00000121297	ok
2	chr6:30554074-30554096	+	4	GTCTGCAT[GGTGAGTGCCGC]	0	E	ABCF1	ENSG00000204574	ok
3	chr11:118793326-118793348	-	3	GGGTACAG[GCTGGGTGCCGC]	2525	I	UPK2	ENSG00000110375	ok
4	chr22:41809060-41809082	+	3	GCGCGCCG[GCTGTGTGCCGC]	13730	-	TEF	ENSG00000167074	ok
5	chr3:32863249-32863271	+	4	GCGGGCCC[GCTGCGTGCCGC]	2825	I	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	GCGCAGAG[GCTGAGTGGCGC]	1956	-	SAMD4A	ENSG00000020577	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, E = exonic, - = 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
α-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 α-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 µl 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 µl 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 genes *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 ± 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.