Stem Cells and Development

Stem Cells and Development: http://mc.manuscriptcentral.com/scd

Simplified footprint-free Cas9/CRISPR editing of cardiac-associated genes in human pluripotent stem cells

Journal:	*Stem Cells and Development*
Manuscript ID	SCD-2017-0268.R1
Manuscript Type:	Original Research Report
Date Submitted by the Author:	05-Feb-2018
Complete List of Authors:	Kondrashov, Alexander; University of Nottingham, Wolfson Centre for Stem Cells, Tissue Engineering & Modelling (STEM) Duc Hoang, Minh ; University of Nottingham, Wolfson Centre for Stem Cells, Tissue Engineering & Modelling (STEM) Smith, James; University of Nottingham, Wolfson Centre for Stem Cells, Tissue Engineering & Modelling (STEM) Bhagwan, Jamie; University of Nottingham, Wolfson Centre for Stem Cells, Tissue Engineering & Modelling (STEM) Duncan, Gary; University of Nottingham, Wolfson Centre for Stem Cells, Tissue Engineering & Modelling (STEM) Mosqueira, Diogo; University of Nottingham, Wolfson Centre for Stem Cells, Tissue Engineering & Modelling (STEM) Munoz, Maria; University of Nottingham, Wolfson Centre for Stem Cells, Tissue Engineering & Modelling (STEM) Vo, Nguyen ; University of Nottingham, Wolfson Centre for Stem Cells, Tissue Engineering & Modelling (STEM) Denning, Chris; University of Nottingham, Wolfson Centre for Stem Cells, Tissue Engineering & Modelling (STEM);
Keyword:	Cardiomyocytes, Differentiation, Embryonic Stem Cells, Genetic Engineering
Manuscript Keywords (Search Terms):	CRISPR/Cas9, cardiomyocytes, Gene editing, Footprint-free
Abstract:	Modelling disease with hPSCs is hindered because the impact on cell phenotype from genetic variability between individuals can be greater than from the pathogenic mutation. While 'footprint-free' Cas9/CRISPR editing solves this issue, existing approaches are inefficient or lengthy. Here, a simplified PiggyBac strategy shortened hPSC editing by 2 weeks and required one round of clonal expansion and genotyping rather than two, with similar efficiencies to the longer conventional process. Success was shown across 4 cardiac-associated loci (ADRB2, GRK5, RYR2, ACTC1) by genomic cleavage and editing efficiencies of 8-93% and 8-67%, respectively, including mono- and/or bi-allelic events. Pluripotency was retained, as was differentiation into high purity cardiomyocytes (CMs; 88-99%). Using the GRK5 isogenic lines as an exemplar, chronic stimulation with the b-adrenoceptor agonist, isoprenaline, reduced beat rate in hPSC-

	CMs expressing GRK5-Q41 but not GRK5-L41; this was reversed by the b-blocker, propranolol. This simplified, footprint-free approach will be useful for mechanistic studies.

Simplified footprint-free Cas9/CRISPR editing of cardiac-associated genes in human pluripotent stem cells

Alexander Kondrashov*†, Minh Duc Hoang†, James G.W. Smith, Jamie R. Bhagwan, Gary Duncan, Diogo Mosqueira, Maria Barbadillo Munoz, Nguyen T.N. Vo, Chris Denning*

Department of Stem Cell Biology, Centre of Biomolecular Sciences, University of Nottingham, NG7 2RD. United Kingdom

† Authors contributed equally to work

* To whom correspondence should be addressed.
Tel: +44(0)115 8231233; Fax: +44(0)115 8231230; Email: a.kondrashov@nottingham.ac.uk
Tel: +44(0)115 8231236; Fax: +44(0)115 8231230; Email: chris.denning@nottingham.ac.uk

Short title: Simplified CRISPR editing in hPSCs

Author email addresses:

Alexander Kondrashov	PAZAVK@exmail.nottingham.ac.uk
Minh Duc Hoang	v.duchm3@vinmec.com
James G.W. Smith	James.Smith@nottingham.ac.uk
Jamie R. Bhagwan	msxjb4@exmail.nottingham.ac.uk
Gary Duncan	gary.duncan91@gmail.com
Diogo Mosqueira	mzxdm@exmail.nottingham.ac.uk
Maria Barbadillo Munoz	mgzmdm@exmail.nottingham.ac.uk
Nguyen T.N. Vo	mzxnnv@exmail.nottingham.ac.uk
Chris Denning	chris.denning@nottingham.ac.uk

Key words:
Cas9/CRISPR; PiggyBac; gene editing; human pluripotent stem cells; genetic disease modelling; cardiomyocytes;

Abstract:

Modelling disease with hPSCs is hindered because the impact on cell phenotype from genetic variability between individuals can be greater than from the pathogenic mutation. While 'footprint-free' Cas9/CRISPR editing solves this issue, existing approaches are inefficient or lengthy. Here, a simplified *PiggyBac* strategy shortened hPSC editing by 2 weeks and required one round of clonal expansion and genotyping rather than two, with similar efficiencies to the longer conventional process. Success was shown across 4 cardiac-associated loci (*ADRB2*, *GRK5*, *RYR2*, *ACTC1*) by genomic cleavage and editing efficiencies of 8-93% and 8-67%, respectively, including mono- and/or bi-allelic events. Pluripotency was retained, as was differentiation into high purity cardiomyocytes (CMs; 88-99%). Using the GRK5 isogenic lines as an exemplar, chronic stimulation with the β-adrenoceptor agonist, isoprenaline, reduced beat rate in hPSC-CMs expressing GRK5-Q41 but not GRK5-L41; this was reversed by the β-blocker, propranolol. This shortened, footprint-free approach will be useful for mechanistic studies.

Introduction

Human pluripotent stem cells (hPSCs) comprise both human embryonic stem cells (hESCs), derived from the inner cell mass of the preimplantation embryo, and human induced pluripotent stem cells (hiPSCs), derived by epigenetic reprogramming of somatic cells [1]. It is now well established that hPSCs are an important modality for biomedicine, with application ranging from understanding human development through to use of their differentiated progeny in safety assessment of drugs, accelerating drug use towards clinic and modelling genetic disease [1]. Suitability in several clinical trials has been, or is being evaluated, including for spinal cord injury, macular degeneration and heart disease [2]. A difficulty that has emerged for the *in vitro* assays is genetic variation between unrelated individuals may cause greater phenotypic differences than do the disease-associated polymorphism(s) [3]. Therefore, creation of isogenic pairs, wherein only the polymorphism of interest differs between lines, is now considered the gold standard. While the number of reports using conventional gene targeting in hPSC is low, the advent of nuclease-mediated targeting, particularly with Cas9/CRISPR, has made precise modification of the genome relatively routine [1].

Despite these advances, difficulties still remain in gene editing of hPSCs. Making single base pair substitutions is technologically challenging when compared to, for example, gene knockouts, where libraries of guide RNAs (gRNAs) are being used in functional genome-wide screens [4]. An important consideration for editing is that, other than the desired polymorphic changes, the level of genome modification post-gene edited hPSC line should be minimal. This is because residual footprints left behind after targeting can alter or abolish neighbouring gene expression [1,5,6]. This advocates the use of footprint-free or scarless approaches.

One route to achieving footprint-free editing is via the delivery of ribonucleoprotein combinations that comprise recombinant Cas9 protein, *in vitro* transcribed gRNA and a ~50-150 base single-stranded DNA oligonucleotide (ssODN) template, which carries the polymorphic change(s) of interest [7]. We demonstrated the utility of this approach by modifying the *ADRB2* locus, which encodes the β2-adrenocetor [1], while others have altered additional loci [7,8]. Although this route is attractive and less toxic than plasmid approach [7], it requires high transfection rates of large complexes, which can be difficult in sensitive cells such as hPSCs. The lack of a drug selection marker also that means considerable screening effort is needed to identify positive clones. An alternative to achieving seamless editing by using ssODNs as a template is via a system termed "*CORRECT*" [9]; however, this requires two sequential clonal selection/expansion steps.

An alternative for footprint-free editing is the *PiggyBac* transposon system [10], although this does require a *TTAA* quadra-nucleotide site for recombination (see Fig. 1). In this approach, a targeting vector

contains a positive-negative drug selection cassette (e.g. Puro-ΔTK; Fig. 1A) that is flanked by *PiggyBac* recombination sites. In turn, these components are flanked by regions of up to 1kb in length that are homologous to the endogenous target locus, thus enabling recombination between template and genome. The desired polymorphism(s) is carried within one arm of homology. Experimentally, the approach is implemented via two sequential steps. First, the targeting vector is co-transfected with plasmids carrying guide RNA and Cas9 to promote genomic cleavage and insertion via homology directed repair into the locus of interest. Survival during positive selection with antibiotics (e.g. puromycin) identifies the hPSC clones that express the cassette, which are then picked, expanded and genotyped (Fig. 1B). Second, antibiotic resistant hPSCs are transfected with a plasmid expressing transposase, which induces internal recombination between *PiggyBac* sites, excision of the selection cassette and reconstitution of a footprint-free locus (Fig. 1B). Colonies that fail to excise the cassette continue to express ΔTK and hence are negatively selected against by the prodrugs, ganciclovir or fialuridin. This leaves the surviving colonies, which can be picked, expanded and genotyped for a second time.

Several reports have described the successful use of this *PiggyBac* approach in hPSC [11,12,13]. Nevertheless, the requirement for two rounds of clonal selection and genotyping over a lengthy timeline is problematic. Particularly for hPSCs, the number of cumulative population doublings correlates genetic [14] and epigenetic [15,16] instability, thereby affecting their downstream applications [17]. Similarly, in mouse iPSCs, genetic instability has been reported within as few as 4-6 passages [18]. Thus, processes that enable gene editing in shorter timelines would be beneficial [19].

In this report we adapted a footprint-free *PiggyBac*-based Cas9/CRISPR gene editing strategy to both simplify and shorten the process. Only one round of clonal selection and genotyping is needed, reducing the process from 49 to 35 days, a 25-30% time saving that equates to ~14 population doublings in hPSCs. We have demonstrated the utility of this simplified approach by making single or dual polymorphic changes to 4 cardiac-related genes, *ADRB2*, *GRK5*, *RYR2* and *ACTC1.* For each of the engineered hPSC lines created, we showed that the cells retained expression of pluripotency markers, a stable karyotype and the ability to differentiate at high efficiency into beating cardiomyocytes that express α-actinin. As an exemplar, we showed significant differences in functional consequence between isogenic pairs of hiPSC-CMs that carry GRK5-L41 or GRK5-Q41 polymorphisms in response to chronic β-adrenergic stimulation and β-blocker rescue. Thus, the approach described provides a simplified and abbreviated route towards mechanistic understanding of how single polymorphic variants alter heart function.

Materials and Methods

Cell culture

All culture was at 37°C at 5% CO_2 in a humidified atmosphere. Unless otherwise stated, all reagents were from ThermoFisher. HUES7 hESCs were gifted by Chad Cowan and Doug Melton at the Harvard Stem Cell Institute. Fibroblasts were derived under ethical consent from individual with the genotypes *RYR2*$^{6739C/T}$ (NRES Committee East Midlands – Nottingham 2 approval 09/H0408/74) and *ACTC1*$^{301G/G}$ (Biomedical Institute of A Coruna, INIBIC). Reprogramming to hiPSCs was via CytoTune 2.0 (ThermoFisher), according to the manufacturer's instructions. Culture was in E8 medium on Matrigel, although processes could also be completed in hESC medium conditioned using mouse embryonic fibroblasts [20]. In the first 4-5 passages after reprogramming, cell harvesting was done using 0.5mM EDTA and thereafter with accutase.

Transfection Optimisation

For transfection and electroporation experiments, hPSCs were seeded at 3x10^5 cells/well of the Matrigel-coated 6 well plate or resuspended cells at 2x10^5 cell/well/transfection condition in Nucleocuvette Strip (16 wells), respectively. Plasmids were transfected into hPSCs using either FuGene HD transfection reagent (Promega, E2311) following the manufacturer's instructions using a ratio between reagent and plasmid DNA of 4:1. To optimise the electroporation using the Amaxa 4D system (Lonza), pmaxGFP plasmid provided in the Lonza Amaxa 4D kit was transfected into hPSCs with human stem cell P3 solution (programs: CA-137, CB-150, CD-1118, CE-118, CM-113, DC-100, DN-100, as recommended by the manufacturer's protocol). The GFP signal was captured using Operetta High-content imaging system (Perkin Elmer) and analysed using Harmony High-content imaging software.

Targeting vector construction

The *ADRB2* targeting vector was constructed via Gibson assembly by using Gibson Assembly master mix (E2611S NEB). Overlapping fragments were produced by PCR (GoTag polymerase, Promega) for three inserts: Dual drug selection cassette (Puro-ΔTK) flanked by *PiggyBac* recombination sites; and the left and right homology regions for *ADRB2* (~1kb upstream and ~1kb downstream of the locus cut site). Primers used are shown in Supplementary Table 1. An EcoRV digested pBluescript backbone plasmid sequence was used as the fourth DNA fragment in the Gibson assembly. A 20µl reaction containing 0.24 pmol of each insert, 0.08pM of Bluescript backbone and 1X Gibson Assembly® Master Mix (NEB) was heated at 50°C for 60 minutes. Subsequent transformation into Top10 competent cells and colony sequencing identified

correctly assembled plasmids. The same approach was used to generate the *GRK5*, *ACTC1* and *RYR2* targeting constructs.

Gene targeting in hPSCs

FuGene HD (Promega) transfection required seeding of $3x10^5$ hPSCs into each well of a Matrigel-coated 6-well plate. Twenty-four hours later, cells were transfected 3.3 µg of CRISPR plasmid components (targeting plasmid, gRNA, Cas9). For Amaxa 4D nucleofection (Lonza), $3x10^6$ hPSCs and 3 µg of CRISPR plasmid components were used with P3 solution, program CA-137. Transfected and nucleofected cells were maintained in E8 medium on Matrigel (hESC medium conditioned using mouse embryonic fibroblasts [20] could also be used). Twenty-four hours post-transfection, medium was supplemented with puromycin (0.25 to 7.5 µg/mL; cell line dependent) for positive selection of clones up to two weeks. The puromycin-positive clones were then harvested and expanded as described in the cell culture section. For cassette excision, cells were seeded at $3x10^5$ cells/well of a Matrigel-coated 6-well plate before delivering transposase plasmid by transfection (3 µg) using FuGene HD transfection as described above. Cells were reseeded to 10 cm dishes, incubated for 2-3 days to allow recombination by transposase and then exposed to medium containing ganciclovir (2 µg/mL) for negative selection of *PiggyBac* excision. Approximately 7-10 days later, clones were manually dissected and genotyped using primers shown in Supplementary Table 1. See this Table and also Figure 3 for location of primers to test for off-target and random integration events. Realtime qPCR to the ampicillin gene was conducted by GoTag® qPCR Master mix (Promega, #A6001) on Applied Biosystems SDS 7500 Fast Real-time PCR template for 45 cycles. Melting curves was obtained for all experimental runs. Relative expression of genes was calculated and expressed as $2^{-\Delta\Delta Ct}$, normalised using *18S*.

Characterisation of hPSC

A) Cardiomyocyte differentiation:

Undifferentiated hPSCs were seeded onto Matrigel-coated dishes at a density of $4x10^4$ cells/cm^2 and allowed to expand for 48h (~80% confluency). At this stage (d1 of differentiation), cultures were treated with medium comprising StemPro34 supplemented with [1:100 dilution) Matrigel and [1 ng/ml] BMP4 (R&D systems). After 24h (d2 of differentiation), medium comprising StemPro34 with [10 ng/ml] BMP4 and [8ng/ml] Activin A (Life Technologies). Medium exchange was performed on d4 of differentiation using RPMI supplemented with 1xB27 (Life Technologies) and small molecule inhibitors, KY02111 (10 µM) and XAV939 (10 µM) (R&D systems). From d8 onwards, cells were maintained in RPMI medium supplemented

with B27 only, with medium changes every 3 days. Cardiac differentiation efficiency was accessed by using immunocytochemistry with primary mouse anti-human α-actinin antibody (Sigma #A7811, 1:800) dilution; secondary goat anti-rabbit Alexa633 (Invitrogen #A21052, 1:400); counterstaining with 0.5 µg/ml DAPI (Sigma #D9542, 1:500). Immunofluorescence images were captured using Operetta High-content system (Perkin Elmer) and analysed using Harmony high-content analysis software.

B) Gene expression:

RNA was isolated from undifferentiated hPSCs and derived cardiomyocytes at day 14 of differentiation using RNeasy mini kit (Qiagen). Synthesis of cDNA was carried out using 1 µg RNA with SuperScript III Reverse Transcriptase kit (Invitrogen), according to manufacturer instructions. *ADRB2* analysis was with Taqman qPCR (Applied Biosystems, #Hs00240532_s1) and signals were normalised to *GAPDH* (Applied Biosystems, #Hs99999905_m1) as the housekeeping gene, following the manufacturer's instructions. Semi-quantitative PCR cycle conditions were 95°C for 2 min, 64.5°C for 30 sec (*GRK5*, *ACTC1*, *RYR2* and *ACTB*) and 72°C for 60 sec, with a final elongation step of 72°C for 10 min. Each reaction used 250 ng of cDNA with Phusion polymerase (NEB) for 35 cycles. Gels were imaged with a LAS-4000 (Fujifilm) image analyser, densitometry was carried out using FIJI, a version of ImageJ (National Institutes of Health) and signals were normalised to *ACTB* as the housekeeping gene. Primers for expression analysis are provided in Supplementary Table 1.

C) Immunocytochemistry analysis of nuclear pluripotent markers:

Human PSCs were cultured at 30,000 cells/cm^2 in Matrigel-coated 96-well plates (Perkin Elmer CellCarrier) until reaching 60% confluent before fixing with 4% PFA. Fixed cells were perforated using 0.01% Triton X-100 and 0.05% Tween 20 (diluted in PBS). The cells were then incubated with mouse-anti human OCT4 (C-10 clone, Santa Cruz Biotech #sc-5279, 1:100) and subsequent secondary antibody using goat-anti mouse Alexa488 (Invitrogen #A11001, 1:1000), counterstained with 0.5 µg/ml DAPI. Immunofluorescence images were captured using Operetta High-content system (Perkin Elmer) and analysed using Harmony high-content analysis software.

D) Flow cytometry analysis of surface pluripotent markers

To analyse surface markers, hPSCs were harvested and fixed using 4% PFA followed by incubation with PE-conjugated SSEA-1 (eBioMC-480 clone, ThermoFisher #12-4752, 1:100), SSEA-4 (eBioMC-813-70 clone, ThermoFisher #12-8843, 1:200) and TRA-1-81 (TRA-1-81 clone, ThermoFisher #12-8883, 1:100)

antibodies for 20 min at 4°C. Cells were analysed using an FC500 Flow cytometer (Beckman Coulter) and data were analysed with FlowJo software.

E) Karyotyping: Metaphase spreads were prepared as previously described [20] from hPSCs after final genotype was confirmed, and karyotype analysis was performed by G-banding of 30 metaphase spreads in each sample, according to guidelines from the International System for Human Cytogenetic Nomenclature.

Functional analysis of GRK5 hPSC-cardiomyocyte polymorphic variants

To measure the beat rate of CMs in real time, the CardioExcyte96 system (Nanion) was used. Briefly, the 96-well sensor plates of the CardioExcyte96 were coated by incubation (1.5 h) with fibronectin at 1:100 dilution in PBS (without Ca^{2+} and Mg^{2+}). CMs at d25 to d28 were dissociated and seeded onto the sensor plate at 60,000 cells/well. Plates were incubated for 48h before changing the medium and starting the recordings according to the following timeline: 0-2 h, baseline recording; 2 h, spike with 100nM isoprenaline; 24 h, repeat spike of isoprenaline; 48-50 h, end of recording. Beat rate of CMs was recorded throughout the experiments at intervals of 2 to 10 minutes. For the non-selective beta-blocker experiment, propranolol (200 nM) was added 1 hour before starting isoprenaline treatment and maintained throughout.

Results

Locus selection and targeting strategy for *ADRB2* (β2-adrenoceptor)

Over the course of multiple experiments in our laboratory, we observed Cas9/CRISPR gene targeting efficiencies of ~30% (158 of 421 colonies assessed) across 12 different loci and/or hPSC lines (data not shown) when using optimised transfection conditions (Supp. Fig. 1). In the context of the 2 step *PiggyBac* process, we reasoned that the gene targeting efficiency during step 1 would be rate limiting because cassette excision should occur in most cells, provided transposase delivery is at high efficiency at the start of step 2. An alternative strategy could be to merge steps 1 and 2 of the *PiggyBac* process. This would have the advantage of not only simplifying editing, but also of reducing the time to produce gene modified hPSCs by 14 days; this equates to ~14 population doublings and 25-30% of the whole targeting process (Fig. 1B).

To test this notion, we selected the *ADRB2* locus for several reasons. *ADRB2* encodes β2-adrenoceptor, a G-protein coupled receptor that has an N-terminal domain positioned in the extracellular compartment. In this domain, two polymorphic variants at amino acid positions p.Gly16Arg (c.*G46A*) and p.Glu27Gln (c.*G79C*) alter patient response during heart failure [21]. Thus, production of isogenic hPSC lines from which cardiomyocytes can be produced would be beneficial in understanding the mechanism of these differences. We also selected this locus because it is expressed in undifferentiated hPSCs, albeit at much lower levels than in hPSC-cardiomyocytes (Fig. 1C). This may be useful since an 'open' configuration is considered to be more permissible to gene targeting [22]. However, *ADRB2* also requires a footprint-free strategy because it is a single exon gene with complex 5' and 3' untranslated regions, which include multiple regulatory elements and domains required for proper expression of *ADRB2* and its membrane targeting [23,24,25]. As such, positioning a selection cassette or a short footprint in these regions may be disruptive to cell signalling and function, even in the undifferentiated state.

The *PiggyBac* approach requires an endogenous quadra-nucleotide *TTAA* palindrome sequence at the site of recombination, which theoretically occurs at 329bp intervals through the genome [26]. However, the *PiggyBac* transposon has a preference for areas surrounding transcription start sites and CpG islands [27], suggesting that even distribution of *TTAA* sites does not occur. Supporting this notion, our analysis of the genomic regions flanking the position 46 or 79 *ADRB2* polymorphic variants in HUES7 hESCs revealed that the nearest *TTAA* site was 748 bases away (data not shown), which far exceeds the distance recommended for insertion via nuclease-mediated targeting [10]. However, we noted the sequence *CTC ATC* (nucleotide position 124-129) situated 45 bases downstream of the position 79 polymorphism in *ADRB2* coding sequence; codon redundancy for leucine meant that substitutions could be made to *TTA ATC*, which created the *TTAA* site necessary for *PiggyBac* recombination whilst being synonymous and retaining the native Leu41-Ile42 peptide sequence (Fig. 2). We sought to minimise any further changes, silent or otherwise, to the *ADRB2* locus. Therefore we selected a gRNA with a protospacer adjacent motif (PAM) overlapping the polymorphic change c.*G79C*, ensuring cleavage of genomic, but not targeting vector, sequences would occur (Fig. 2; Supp Table 1). Thus, the left arm of homology in the targeting vector contained c.*G46A* (p.Gly16Arg), c.*G79C* (p.Glu27Gln) and c.*C124T*/c.*C126A* (synonymous: p.Leu41-Ile42) modifications directed towards the *ADRB2* locus (Fig. 1A; Fig. 2).

Simplified *PiggyBac* gene editing in *ADRB2* in hPSCs

The process outlined in Fig. 1 entails 2 steps, with gene targeted insertion of the *PiggyBac* cassette and the associated polymorphic changes occurring in the first step, followed by transposase-mediated cassette removal in the second step. Since we anticipated cassette excision should occur at high efficiency

(Supp Fig. 1), we wished to test whether the frequency and types of targeting events was similar after first (midpoint) and second (end) steps. In addition, we wanted to ensure that streamlining the process by progressing directly from positive (puromycin) to negative (ganciclovir) selection did not have a detrimental effect.

HUES7 hESCs were co-transfected with Cas9, gRNA and *ADRB2* targeting plasmids, and then subjected to puromycin treatment. Once early stage drug resistant colonies had formed, a portion of the colonies were picked for genotyping after step 1. The remainder of the cells were harvested, transfected with transposase and then treated with ganciclovir, before allowing colonies to form for picking and genotyping after step 2. All clones were assessed by PCR amplification coupled to direct sequencing across the left arm of homology (Fig. 3A; Supp Table 1; Supp. Fig. 2A,B). Genotyping after first vs second step showed high frequencies (Fig. 3B), wherein genomic cleavage was evident in 8/11 (73%) and 6/12 (50%). Specifically between categories 18% vs 8% untargeted, 9% vs 8% mono-allelic targeting, 9% vs 33% bi-allelic targeting, 55% vs 8% indels, indicated by messy reads around Cas9 cleavage site, and 9% vs 42% unclear result, indicated by PCR failure or lack of sequencing data (Fig. 3A; Supp. Fig 2A,B).

We also evaluated off target events (Fig. 3C; Supp Table 1). We focussed on known coding or regulatory sequences where gRNAs had full PAM site complementarity and/or fewer than 5 mismatches with the target. The 5 putative sites that met these criteria were shown by PCR amplification and sequencing to be unaffected by off targeting (Fig. 3C). Therefore, the simplified *PiggyBac* approach was successfully used to produce an isogenic set of wildtype (untargeted), heterozygote (mono-allelic) and homozygote (bi-allelic) dual-site modifications at nucleotide positions 46 and 79 in the 5' end of the *ADRB2* gene in hESCs.

Finally, we tested for unwanted random integration events of the vector elsewhere in the genome by PCR (Fig. 3D). As expected, control primers that spanned the PAM site in *ADRB2* gave a product from parental cells and after step 2, but not step 1 since the bi-allelic presence of a complex puro-ΔTK cassette blocks the PCR reaction. Correspondingly, PCR products specific to the *ADRB2-puro-ΔTK* junction and to *ΔTK* were produced only from step 1 samples, indicating that no residual targeting selection cassette could be detected after transposase-mediated removal. Remnants of the pBlueScript plasmids backbone were tested for by PCR to the ampicillin gene. No products were seen by conventional PCR (data not shown). Therefore, qPCR was carried out using a positive control, wherein targeting plasmid DNA was diluted to the equivalent of a single genomic copy into parental HUES7 DNA. Relative to this positive control, samples from parental cells, step 1 and step 2 gave a signal 10- to 20-fold lower. Collectively, these data suggest that precise targeting of the selection cassette occurred only at the *ADRB2* locus and not at random elsewhere in the genome, and cassette excision occurs after transposase-mediated removal.

Applying simplified *PiggyBac* gene editing to other cardiac-associated loci in hPSCs

Efficiency of gene targeting, including using Cas9/CRISPR, is known to be influenced by genomic environment, including complexity and GC-richness of gene sequence, active gene expression, availability of sites to guide nuclease docking, and cell type. Therefore, we selected 3 additional cardiac-associated loci with different genetic properties but each with relevance to human health or heart disease (Fig. 4).

GRK5 encodes GPCR specific kinase involved in β-adrenergic receptor desensitisation. It has been suggested that a c.*A122T* (p.Gln41Leu) polymorphism causes a natural β-blocker effect that may be protective against heart disease [28]. *ACTC1* encodes cardiac actin and a mutation at c.*G301A* (p.Glu101Lys) causes hypertrophic cardiomyopathy, altered calcium sensitivity, arrhythmias and, in some cases, sudden cardiac death [29]. Finally, *RYR2* encodes ryanodine receptor, which is a calcium release channel in the sarcoplasmic reticulum. A highly malignant mutation of c.*C6737T* (p.Ser2246Leu) causes catecholaminergic polymorphic ventricular tachycardia (CPVT), which can lead to sudden cardiac death [30].

All 4 genes were expressed in undifferentiated hPSCs (Figs 1 & 5), which is surprising since *ACTC1* encodes for cardiac actin, a cardiomyocyte specific structural protein (Fig. 5). The GC content of the region surrounding the polymorphisms, gRNA and *TTAA* sites differs between *ADRB2* (64%), *GRK5* (56%), *ACTC1* (53%) and *RYR2* (42%) (Figs. 2 & 5). Thus, this set provided an opportunity to test the simplified *PiggyBac* approach in genes differing in sequence composition and that were expressed at relatively low levels in hPSCs.

Each gene was targeted in a different hPSC line out of necessity. The starting genotypes were hESC (line HUES7) $GRK5^{122A/A}$, hiPSC $ACTC1^{301G/G}$ from a healthy individual within a family with familial hypertrophic cardiomyopathy and hiPSC $RYR2^{6737C/T}$ from a young patient with CPVT. In designing the targeting strategies (Fig. 5), we elected to use endogenous *TTAA* sites for *PiggyBac* recombination that resided in neighbouring introns. In addition, for each of the three genes (*GRK5*, *ACTC1* and *RYR2*), gRNAs were chosen that spanned these *TTAA* sites; this means that the gRNAs recognised the endogenous genomic sequence but not the targeting vector because the *TTAA* demarcates the *PiggyBac* cassette insertion site (Fig. 5).

Adopting these two strategies allowed production of true isogenic lines; i.e. no further sequence changes with potentially unknown effects were required either to form a *de novo TTAA* site or to protect the targeting vector from gRNA/Cas9 cleavage. The potential disadvantage of this approach is that the distance between gRNA/Cas9 cleavage site and the desired polymorphic change is increased, which raises the likelihood of recombination occurring between these two locations and hence not carrying the polymorphic change into the genome. Indeed, while the distance between PAM site and polymorphic

change was 107bp and 136bp for *GRK5*$^{122A/A}$ and *RYR2*$^{6737C/T}$ respectively, it was 313bp for *ACTC1*$^{301G/G}$ (Fig. 4).

Targeting vectors were constructed for these genes (Figs. 4 & 5A) using the same design principles that were used for *ADRB2* and thus relied on ~2kb of total homology, with ~1kb in each of the left and right arms (Fig. 1). Following positive (puromycin) and then negative (ganciclovir) selection, colonies were expanded for PCR and sequence analysis (Fig. 5C). For all three genes, successful targeting of the polymorphisms to the left arms was observed with concurrent excision of the *PiggyBac* selection cassette and reconstitution of the endogenous *TTAA* site (Fig. 5C,D). However, the targeting efficiencies differed considerably (Fig. 5D). In *GRK5*, genomic cleavage was confirmed in 93% clones, of which 13% and 47% were mono-allelic and bi-allelic targeting events respectively. This overall trend of correct targeting was similar to *ACTC1*, where cleavage was 75%, although this led to 67% and 0% mono-allelic and bi-allelic targeting events, respectively. In contrast, cleavage was only evident in 8% of *RYR2* clones, which converted to a successful editing event. In summary, the simplified approach was used to produce footprint-free, isogenic pairs for 4 cardiac-related genes in hPSCs.

Characterisation of gene edited hPSCs

Although correct targeting had been achieved, it was important to confirm whether specific pluripotency and differentiation characteristics were retained in *ADRB2*, *GRK5*, *ACTC1* and *RYR2* gene edited hPSC lines. Representative examples are shown (Fig. 6) but similar results were obtained from multiple clones, with the exception of *RYR2* where only one successful targeting event was identified. In all cases, immunostaining coupled with high content image analysis showed that almost all cells expressed the pluripotency marker, OCT4. This was supported by flow cytometry, where 78-99% and 76-100% of hPSCs being positive for TRA-1-81 and SSEA4, whereas <3% displayed the differentiation marker, SSEA1 (Figure 6).

The metaphase spreads of 30 cells per line were assessed by G-banding karyotyping. Assembly of homologous chromosomes into a karyogram showed no evidence of aberration. Finally, directed monolayer differentiation was used on each line to induce beating sheets of cardiomyocytes. These were dispersed on day 12-15 of differentiation and stained with α-actinin, before using high content image analysis to show cardiomyocyte purity was between 88 and 98%. Thus, the edited lines retained key characteristics of pluripotency, most notably differentiation to functional cardiomyocytes.

Evaluating consequences of GRK5-L41 and –Q41 variants on hPSC-CM function

To demonstrate the utility of isogenic sets of hPSC lines, we selected wild type GRK5$^{122A/A}$ and homozygote edited GRK5$^{122T/T}$ lines, which differ only in leucine (L) or glutamine (Q) at position 41 of the encoded peptide. It has been suggested that the GRK5-L41 variant acts as a natural β-blocker and so is protective against adrenergic stress in the heart [28]. Therefore, we seeded confluent monolayers of cardiomyocytes derived from the GRK5 isogenic lines onto the CardioExcyte-96 impedance platform to assess beating characteristics during chronic (up to 50 hours) stimulation with the β-adrenoceptor agonist, isoprenaline (Fig. 7).

During the first 30 hours of isoprenaline treatment, cardiomyocytes from both variants showed similar responses with maximum beat rates reaching ~150% of baseline values (Fig. 7Ai, Aii). This similarity was confirmed by calculating normalised beat rate (GRK5-Q41 divided by GRK5-L41), which gave values of close to 1 (Aiii). However, from 30 hours onwards, the normalised rate of GRK5-Q41 declined, finally reaching 60-80% of baseline by the 38-48 time window. In contrast, by the end of the evaluation period, GRK5-L41 maintained an average rate of 150%, which was reflected in a Q41/L41 response ratio of ~0.5 (note arrows in Fig. 7Ai, ii, iii). This mirrors *in vivo* findings, which show that, unless compensation mechanisms can be invoked, prolonged (>30h) activation of adrenoceptors by catecholamines compromises cardiomyocyte recovery [28].

Since the GRK5-L41 variant has been suggested to impart a mild protective effect during chronic β-adrenergic stimulation, we re-ran the experiment but this time with co-incubation of isoprenaline and the non-specific β-blocker, propranolol (Fig. 7B). As expected, the initial chronotropic response of both variants was subdued by propranolol. Notably, however, the chronic decline in beat rate to well below baseline levels seen by 38-48 hours in GRK5-Q41 with isoprenaline alone (Fig/. 7Bi) was abolished with the addition of propranolol (Fig. 7Bii) and was reflected by response rate ratios of close to 1 throughout the timecourse (Fig. 7Biii). Thus, chronic overstimulation of the β-adrenoceptor system eventually caused a decline in beat rate in GRK5-Q41, but not GRK5-L41 hPSC-CMs, and this could be reversed by β-blockade. This provides a tool for mechanistic understanding of genotype-phenotype interactions, which we are now investigating.

Discussion

We successfully demonstrated a simplified footprint-free approach to gene edit 4 distinct cardiac-associated loci in hPSCs, with modifications including mono- and/or bi-allelic targeting. This included introducing polymorphic changes in hESC and/or hiPSC lines that were anticipated to be mildly beneficial to cardiomyocyte function into *ADRB2* and *GRK5* or severely damaging into *ACTC1*. We also corrected a damaging mutation in the *RYR2* gene. The edited hPSC lines retained the ability to undergo high efficiency differentiation to cardiomyocytes, enabling us to demonstrate the utility of this approach by showing

functional differences in drug response for the GRK5 isogenic set. This simplified *PiggyBac* approach is easily adaptable to other loci, providing there is appropriate proximity of *TTAA* sites, either native or modified by engineering. Applicability will be irrespective of whether the targeting strategy employs conventional or nuclease (e.g. zinc fingers, TALE, Cas9/CRISPR) strategies and will be of future value in facilitating mechanistic studies.

The need for isogenic hPSC lines was highlighted recently by Sala et al., 2016 [3]. Comparison of action potential duration 90 (APD_{90}), an electrophysiology parameter, in cardiomyocytes derived from 18 hPSC lines showed more than a 4-fold difference, with values ranging from ~140ms to 600ms. Even between different commercial suppliers of hPSC-cardiomyocytes, where quality control is high before release to customers, the range was 225ms to 600ms. A notable departure from this variation was one isogenic pair, where cardiomyocytes from both lines had highly similar APD_{90} values of ~230ms.

Contextually, the normal range for humans APD_{90} values (usually cited as QT interval) is 350-450ms and increases of 10-20% are worrisome. During drug development such prolongation would likely lead to the termination the drug [31]. Clinically, QT intervals of >460-500ms usually signify disease state, such as long QT syndrome, which is caused by mutations in various ion channel proteins and can lead to sudden cardiac death [32]. This means that depending on the hPSCs selected, the phenotypic variation between lines (up to 400%) can be greater than any change caused by the mutation (usually 10 to 100%). This may explain some of the discrepancies reported in the literature for hPSC-based disease modelling, including for the magnitude of change caused by mutations in *KNCQ1*, which underlies long QT syndrome type 1 [33,34]. Consequently, the use of isogenic pairs is becoming the gold standard for disease modelling using hPSCs. The isogenic approach allows desired polymorphisms to be studied within the same genetic background and the 'noise' is eliminated from the other estimated ~11 million SNPs, 2.8 million short indels and ~500,000 block substitutions that exist between unrelated individuals [35].

A true isogenic pair will differ only in the desired polymorphic change. Part or whole remnants of selection cassettes can perturb gene function [36], even when positioned in introns because of the presence of currently unannotated sequences. Indeed, in hPSCs, we found that even when Cas9/CRISPR was used to target Ef1α-driven blasticidin or puromycin resistance markers into neighbouring introns, this abolished expression of *KCNH2 [1]* and *MYH7* (Supp Fig. 3C) genes, which encode the HERG potassium ion channel and beta myosin heavy chain structural protein, respectively. In both cases cassette removal restored expression of *KCNH2* and *MYH7*. For *ADRB2*, the complexity of the locus and absence of a nearby *TTAA* site necessitated conversion of *CTC ATC* to *TTA ATC*. In humans, both *CTC* and *TTA* are compatible with the leucine tRNA machinery but the probability of use is 0.2 and 0.07, meaning that *CTC* is preferred. Also, current gene annotation shows this change should not interfere with control regions (promoters, enhancers, non-coding RNAs, splice sites etc) but needs to be borne in mind during targeting design. Thus,

any changes, from single bases through to residual sequences ([1,37,38]; Supp Fig. 3) may need thorough investigation to rule out any potential negative impact on cell function.

For loci that are more refractory to targeting, cultures can be pooled at the midpoint of the process (after step 1 / puromycin treatment) and an aliquot of cells taken for bulk PCR analysis. Primers are chosen to span from the selection cassette to the flanking genomic region of the locus of interest. If no PCR product is produced this may suggest the experiment should be abandoned. However, if there is a product then the cells can be re-seeded, transfected with transposase and then treated with ganciclovir to finish the excision / colony selection process.

A surprising observation was that when cells at this puromycin resistant midpoint were cryopreserved, the positive-negative selection cassette was silenced upon thawing of the cells; this occurred across several loci beyond those described in this report. We are unsure as to why the cryopreservation-thaw cycle caused this effect. Indeed, it is well documented that silencing of transgenes occurs readily in hPSCs, particularly when non-mammalian promoters are used [39]. However, we used the mammalian promoter, phosphoglycerate kinase (PGK), which is usually well-tolerated [40,41]. We are not aware of other reports where transgene expression is maintained during long-term culture unless a cryopreservation-thaw cycle is introduced.

Although all loci were targeted successfully, there were notable differences. Genome editing occurred at an efficiency of 42-67% in *ADRB2*, *GRK5* and *ACTC1*, but only 8% in *RYR2*. All the genes were expressed but this is not a prerequisite for Cas9/CRISPR targeting. Our data for *MYH7* showed a frequency of mono- and bi-allelic events totalled ~25% (Supp Fig. 3). In terms of GC content, *RYR2* had the lowest (42%) around the target site, which might be expected to give better access for gene targeting rather than the lowly 8% reported here. This may be because the complexity of the *RYR2* locus is high, with regions flanking the target site including repetitive elements (LINE, SINE, Alu). Another parameter that could influence targeting efficiency is the cell line used. Out of necessity we used different hPSC lines because of their starting genotype, which in some cases was disease- or patient-specific. Many similarities and differences have been reported between hPSC lines [16]. In our report, we found that the puromycin concentration required during selection varied from 0.25 to 7.5 µg/mL). Thus, it would be unsurprising if variation extended to differential targeting efficiencies between hPSC lines.

Vector construction and lengths of homology regions are also factors known to impact targeted recombination [42]. The same design principles were used for all 4 loci but the distances between the PAM site in the gRNA and polymorphism (termed PAM-SNP) varied out of necessity. Differences in targeting frequency may be explained by the mechanism of repair. DNA repair occurs via multiple pathways or sub-pathways including DNA double-strand break repair (DSBR), Holliday junction dissolution, synthesis-

dependent strand annealing (SDSA) and single-strand DNA incorporation (ssDI) [43,44]. With regards DSBR, long conversion tracts (approx. ±1 kb) are generated either side of the of conversion zone, with probability of conversion decreasing as a function of PAM-SNP distance [45]. Linear dependency also occurs with ssODNs [9,43], but creates conversion tracts of approx. ±60 nucleotides [43], which is why this approach tends to only incorporate small insertions or substitutions.

The PAM-SNP frequency-distance relationship may explain some of the differences in nature and efficiency of targeting events. For *ACTC1*, with a 300 nucleotide distance, there was a higher probability of recombination occurring between PAM site and polymorphism. After transposase-mediated cassette excision, the sequence in the final chromatogram would appear as untargeted because the approach was designed to be footprint-free. This may have contributed to a profile of clones being untargeted = high (25%), mono-allelic targeted = high (67%) and bi-allelic targeted = low (0%). In contrast, the short PAM-SNP distance of around 100 nucleotides or less for *ADRB2* and *GRK5* presented profiles of 0%, 8%, 33% and 8%, 13%, 47% respectively. Fortuitously, only heterozygote mutations occur in humans for *ACTC1*, presumably because it would likely lead to early lethality, which is the case in mouse knockouts. However, the PAM-SNP distance is clearly not the only factor, since most (92%) clones for *RYR2* were not targeted. We cannot be sure whether the one *RYR2* clone was mono- or bi-allelic allelic targeting event since the template was identical to the healthy allele so only correction of the mutant allele could be detected.

Our main goal in this work was to reduce the duration required to produce isogenic sets of hPSCs, with a specific emphasis on *in vitro* disease modelling of the cardiovascular system. While others have used the *PiggyBac* system, we describe an abbreviated version that not only saves time and effort but the number of population doublings required to produce the gene edited cells. This is important because both empirical experimentation [14] and mathematical modelling [46] shows that genetic and epigenetic change are inevitable as a function of time.

The targeted clones in this study were examined by karyotyping of at least 30 metaphase spreads. Nevertheless, further detailed analysis will be need to examine the broader stability of these lines. The rate of epigenetic change is highest soon after hESC line derivation, with most changes being haphazard [14,15]. In contrast, many genetic changes are predictable. This is exemplified by a large-scale study [14] of 136 hESC and hiPSC lines from 38 laboratories worldwide, which showed a progressive tendency to acquire changes on prolonged culture. Common changes at the chromosome level were part or whole gains of 1, 12 and/or 17. However, in approximately 20% of lines studied, there was also gain of a minimal amplicon in chromosome 20q11.21. This included three genes, *ID1*, *BCL2L1* and *HM13*, with *BCL2L1* driving a selective advantage for hPSC survival in culture. Whether stochastic or non-stochastic, these changes may affect the quality of the cells for biomedical application. Strategies to reduce the population doublings required during their manipulation should be welcomed, though to date this has not been considered. This would

bring genetically engineered hPSCs into kilter with the international guidelines for clinical grade lines, where low passage seed stocks or master banks are recommended [47].

Acknowledgements

This work was supported by (all to CD) the British Heart Foundation [SP/15/9/31605., RG/15/6/31436., PG/14/59/31000., RG/14/1/30588., P47352/Centre for Regenerative Medicine]; BIRAX [04BX14CDLG]; Medical Research Council [MR/M017354/1]; The National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs) [CRACK-IT. FULL PROPOSAL code 35911-259146., NC/K000225/1]; Heart Research UK [TRP01/12]. Funding for open access charge: British Heart Foundation. Plasmids for *PiggyBac* targeting were made available via the Wellcome Trust Sanger Institute. This included pCMV-hyPBase (pcDNA3-based expression vector of a hyperactive piggyBac transposase; [48] and pMCS-AAT-PB:PGKpuro-ΔTK (A donor template vector for genetic correction. The vector contains the puro-ΔTK cassette flanked by the piggyBac repeats; [11].

Author disclosure:

Alexander Kondrashov	No competing financial interests exist
Minh Duc Hoang	No competing financial interests exist
James G.W. Smith	No competing financial interests exist
Jamie R. Bhagwan	No competing financial interests exist
Gary Duncan	No competing financial interests exist
Diogo Mosqueira	No competing financial interests exist
Maria Barbadillo Munoz	No competing financial interests exist
Nguyen T.N. Vo	No competing financial interests exist
Chris Denning	No competing financial interests exist

REFERENCES

1. Denning C, V Borgdorff, J Crutchley, KSA Firth, V George, S Kalra, A Kondrashov, MD Hoang, D Mosqueira, A Patel, L Prodanov, D Rajamohan, WC Skarnes, JGW Smith and LE Young. (2016). Cardiomyocytes from human pluripotent stem cells: From laboratory curiosity to industrial biomedical platform. Biochimica Et Biophysica Acta-Molecular Cell Research 1863:1728-1748.
2. Smith JG, AD Celiz, AK Patel, RD Short, MR Alexander and C Denning. (2015). Scaling human pluripotent stem cell expansion and differentiation: are cell factories becoming a reality? Regen Med 10:925-30.
3. Sala L, M Bellin and CL Mummery. (2016). Integrating cardiomyocytes from human pluripotent stem cells in safety pharmacology: has the time come? Br J Pharmacol.
4. Kim HS, K Lee, S Bae, J Park, CK Lee, M Kim, E Kim, S Kim, C Kim and JS Kim. (2017). CRISPR/Cas9-mediated gene knockout screens and target identification via whole-genome sequencing uncover host genes required for picornavirus infection. J Biol Chem 292:10664-10671.
5. Meier ID, C Bernreuther, T Tilling, J Neidhardt, YW Wong, C Schulze, T Streichert and M Schachner. (2010). Short DNA sequences inserted for gene targeting can accidentally interfere with off-target gene expression. FASEB J 24:1714-24.
6. Pan Y, L Zhang, Q Liu, Y Li, H Guo, Y Peng, H Peng, B Tang, Z Hu, J Zhao, K Xia and JD Li. (2016). Insertion of a knockout-first cassette in Ampd1 gene leads to neonatal death by disruption of neighboring genes expression. Sci Rep 6:35970.
7. Kim S, D Kim, SW Cho, J Kim and JS Kim. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Research 24:1012-1019.
8. DeWitt MA, W Magis, NL Bray, T Wang, JR Berman, F Urbinati, SJ Heo, T Mitros, DP Muñoz, D Boffelli, DB Kohn, MC Walters, D Carroll, DI Martin and JE Corn. (2016). Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. Sci Transl Med 8:360ra134.
9. Paquet D, Kwart D, Chen A, Sproul A, Jacob S, Teo S, Olsen KM, Gregg A, Noggle S, Tessier-Lavigne M. (2016), Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. Nature. 533(7601):125-9.
10. Yusa K. (2013). Seamless genome editing in human pluripotent stem cells using custom endonuclease-based gene targeting and the piggyBac transposon. Nature Protocols 8:2061-2078.
11. Yusa K, ST Rashid, H Strick-Marchand, I Varela, PQ Liu, DE Paschon, E Miranda, A Ordóñez, NR Hannan, FJ Rouhani, S Darche, G Alexander, SJ Marciniak, N Fusaki, M Hasegawa, MC Holmes, JP Di Santo, DA Lomas, A Bradley and L Vallier. (2011). Targeted gene correction of α1-antitrypsin deficiency in induced pluripotent stem cells. Nature 478:391-4.

12. Sun N and HM Zhao. (2014). Seamless Correction of the Sickle Cell Disease Mutation of the HBB Gene in Human Induced Pluripotent Stem Cells Using TALENs. Biotechnology and Bioengineering 111:1048-1053.
13. Xie F, L Ye, JC Chang, AI Beyer, JM Wang, MO Muench and YW Kan. (2014). Seamless gene correction of beta-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. Genome Research 24:1526-1533.
14. Amps K and PW Andrews and G Anyfantis and L Armstrong and S Avery and H Baharvand and J Baker and D Baker and MB Munoz and S Beil and N Benvenisty and D Ben-Yosef and JC Biancotti and A Bosman and RM Brena and D Brison and G Caisander and MV Camarasa and J Chen and E Chiao and YM Choi and AB Choo and D Collins and A Colman and JM Crook and GQ Daley and A Dalton and PA De Sousa and C Denning and J Downie and P Dvorak and KD Montgomery and A Feki and A Ford and V Fox and AM Fraga and T Frumkin and L Ge and PJ Gokhale and T Golan-Lev and H Gourabi and M Gropp and G Lu and A Hampl and K Harron and L Healy and W Herath and F Holm and O Hovatta and J Hyllner and MS Inamdar and AK Irwanto and T Ishii and M Jaconi and Y Jin and S Kimber and S Kiselev and BB Knowles and O Kopper and V Kukharenko and A Kuliev and MA Lagarkova and PW Laird and M Lako and AL Laslett and N Lavon and DR Lee and JE Lee and C Li and LS Lim and TE Ludwig and Y Ma and E Maltby and I Mateizel and Y Mayshar and M Mileikovsky and SL Minger and T Miyazaki and SY Moon and H Moore and C Mummery and A Nagy and N Nakatsuji and K Narwani and SK Oh and C Olson and T Otonkoski and F Pan and IH Park and S Pells and MF Pera and LV Pereira and O Qi and GS Raj and B Reubinoff and A Robins and P Robson and J Rossant and GH Salekdeh and TC Schulz and K Sermon and J Sheik Mohamed and H Shen and E Sherrer and K Sidhu and S Sivarajah and H Skottman and C Spits and GN Stacey and R Strehl and N Strelchenko and H Suemori and B Sun and R Suuronen and K Takahashi and T Tuuri and P Venu and Y Verlinsky and D Ward-van Oostwaard and DJ Weisenberger and Y Wu and S Yamanaka and L Young and Q Zhou and ISC Initiative. (2011). Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. Nat Biotechnol 29:1132-44.
15. Allegrucci C, YZ Wu, A Thurston, CN Denning, H Priddle, CL Mummery, D Ward-van Oostwaard, PW Andrews, M Stojkovic, N Smith, T Parkin, ME Jones, G Warren, L Yu, RM Brena, C Plass and LE Young. (2007). Restriction landmark genome scanning identifies culture-induced DNA methylation instability in the human embryonic stem cell epigenome. Hum Mol Genet 16:1253-68.
16. Allegrucci C and LE Young. (2007). Differences between human embryonic stem cell lines. Human Reproduction Update 13:103-120.
17. Bosman A, A Letourneau, L Sartiani, M Del Lungo, F Ronzoni, R Kuziakiv, V Tohonen, M Zucchelli, F Santoni, M Guipponi, B Dumevska, O Hovatta, SE Antonarakis and ME Jaconi. (2015). Perturbations of

heart development and function in cardiomyocytes from human embryonic stem cells with trisomy 21. Stem Cells 33:1434-46.

18. Liu P, A Kaplan, B Yuan, JH Hanna, JR Lupski and O Reiner. (2014). Passage number is a major contributor to genomic structural variations in mouse iPSCs. Stem Cells 32:2657-67.
19. Wang G, L Yang, D Grishin, X Rios, LY Ye, Y Hu, K Li, D Zhang, GM Church and WT Pu. (2017). Efficient, footprint-free human iPSC genome editing by consolidation of Cas9/CRISPR and piggyBac technologies. Nat Protoc 12:88-103.
20. Burridge PW, D Anderson, H Priddle, MDB Munoz, S Chamberlain, C Allegrucci, LE Young and C Denning. (2007). Improved human embryonic stem cell embryoid body homogeneity and cardiomyocyte differentiation from a novel V-96 plate aggregation system highlights interline variability. Stem Cells 25:929-938.
21. Sotoodehnia N, DS Siscovick, M Vatta, BM Psaty, RP Tracy, JA Towbin, RN Lemaitre, TD Rea, JP Durda, JM Chang, TS Lumley, LH Kuller, GL Burke and SR Heckbert. (2006). Beta2-adrenergic receptor genetic variants and risk of sudden cardiac death. Circulation 113:1842-8.
22. Daer RM, JP Cutts, DA Brafman and KA Haynes. (2017). The Impact of Chromatin Dynamics on Cas9-Mediated Genome Editing in Human Cells. ACS Synth Biol 6:428-438.
23. Drysdale CM, DW McGraw, CB Stack, JC Stephens, RS Judson, K Nandabalan, K Arnold, G Ruano and SB Liggett. (2000). Complex promoter and coding region beta 2-adrenergic receptor haplotypes alter receptor expression and predict in vivo responsiveness. Proc Natl Acad Sci U S A 97:10483-8.
24. Parola AL and BK Kobilka. (1994). The peptide product of a 5' leader cistron in the beta 2 adrenergic receptor mRNA inhibits receptor synthesis. J Biol Chem 269:4497-505.
25. Tholanikunnel BG, K Joseph, K Kandasamy, A Baldys, JR Raymond, LM Luttrell, PJ McDermott and DJ Fernandes. (2010). Novel mechanisms in the regulation of G protein-coupled receptor trafficking to the plasma membrane. J Biol Chem 285:33816-25.
26. Karlin S1, Burge C, Campbell AM. (1992). Statistical analyses of counts and distributions of restriction sites in DNA sequences. Nucleic Acids Res. 20:1363-70
27. Galvan DL1, Nakazawa Y, Kaja A, Kettlun C, Cooper LJ, Rooney CM, Wilson MH. (2009). Genome-wide mapping of PiggyBac transposon integrations in primary human T cells. J Immunother. 32(8):837-44.
28. Liggett SB, S Cresci, RJ Kelly, FM Syed, SJ Matkovich, HS Hahn, A Diwan, JS Martini, L Sparks, RR Parekh, JA Spertus, WJ Koch, SLR Kardia and GW Dorn. (2008). A GRK5 polymorphism that inhibits beta-adrenergic receptor signaling is protective in heart failure. Nature Medicine 14:510-517.
29. Song W, E Dyer, DJ Stuckey, O Copeland, MC Leung, C Bayliss, A Messer, R Wilkinson, JL Tremoleda, MD Schneider, SE Harding, CS Redwood, K Clarke, K Nowak, L Monserrat, D Wells and SB Marston. (2011). Molecular mechanism of the E99K mutation in cardiac actin (ACTC Gene) that causes apical hypertrophy in man and mouse. J Biol Chem 286:27582-93.

30. Barbanti C, A Maltret and D Sidi. (2017). A focus on pharmacological management of catecholaminergic polymorphic ventricular tachycardia. Mini Rev Med Chem.
31. Braam SR, L Tertoolen, S Casini, E Matsa, HR Lu, A Teisman, R Passier, C Denning, DJ Gallacher, R Towart and CL Mummery. (2013). Repolarization reserve determines drug responses in human pluripotent stem cell derived cardiomyocytes. Stem Cell Res 10:48-56.
32. Morita H, J Wu and DP Zipes. (2008). The QT syndromes: long and short. Lancet 372:750-63.
33. Christ T, A Horvath and T Eschenhagen. (2015). LQT1-phenotypes in hiPSC: Are we measuring the right thing? Proc Natl Acad Sci U S A 112:E1968.
34. Greber B, AO Verkerk, G Seebohm, CL Mummery and M Bellin. (2015). Reply to Christ et al.: LQT1 and JLNS phenotypes in hiPSC-derived cardiomyocytes are due to KCNQ1 mutations. Proc Natl Acad Sci U S A 112:E1969.
35. Shen H, J Li, J Zhang, C Xu, Y Jiang, Z Wu, F Zhao, L Liao, J Chen, Y Lin, Q Tian, CJ Papasian and HW Deng. (2013). Comprehensive characterization of human genome variation by high coverage whole-genome sequencing of forty four Caucasians. PLoS One 8:e59494.
36. Friedel RH, W Wurst, B Wefers and R Kühn. (2011). Generating conditional knockout mice. Methods Mol Biol 693:205-31.
37. Bellin M, S Casini, RP Davis, C D'Aniello, J Haas, D Ward-van Oostwaard, LG Tertoolen, CB Jung, DA Elliott, A Welling, KL Laugwitz, A Moretti and CL Mummery. (2013). Isogenic human pluripotent stem cell pairs reveal the role of a KCNH2 mutation in long-QT syndrome. EMBO J 32:3161-75.
38. Zhang M, C D'Aniello, AO Verkerk, E Wrobel, S Frank, D Ward-van Oostwaard, I Piccini, C Freund, J Rao, G Seebohm, DE Atsma, E Schulze-Bahr, CL Mummery, B Greber and M Bellin. (2014). Recessive cardiac phenotypes in induced pluripotent stem cell models of Jervell and Lange-Nielsen syndrome: disease mechanisms and pharmacological rescue. Proc Natl Acad Sci U S A 111:E5383-92.
39. Liew CG, JS Draper, J Walsh, H Moore and PW Andrews. (2007). Transient and stable transgene expression in human embryonic stem cells. Stem Cells 25:1521-8.
40. Anderson D, T Self, IR Mellor, G Goh, SJ Hill and C Denning. (2007). Transgenic enrichment of cardiomyocytes from human embryonic stem cells. Mol Ther 15:2027-36.
41. Norrman K, Y Fischer, B Bonnamy, F Wolfhagen Sand, P Ravassard and H Semb. (2010). Quantitative comparison of constitutive promoters in human ES cells. PLoS One 5:e12413.
42. Deng C and MR Capecchi. (1992). Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. Mol Cell Biol 12:3365-71.
43. Kan Y, Ruis B, Takasugi T, Hendrickson EA. (2017). Mechanisms of precise genome editing using oligonucleotide donors. Genome Res. 2017 Jul;27(7):1099-1111.
44. Davis L, Maizels N. (2016). Two Distinct Pathways Support Gene Correction by Single-Stranded Donors at DNA Nicks. Cell Rep. 2016 Nov 8;17(7):1872-1881.

45. Kan Y, Ruis B, Lin S, Hendrickson EA. (2014). The mechanism of gene targeting in human somatic cells. PLoS Genet. 2014 Apr 3;10(4):e1004251.
46. Olariu V, NJ Harrison, D Coca, PJ Gokhale, D Baker, S Billings, V Kadirkamanathan and PW Andrews. (2010). Modeling the evolution of culture-adapted human embryonic stem cells. Stem Cell Res 4:50-6.
47. Andrews PW, Baker D, Benvinisty N, Miranda B, Bruce K, Brüstle O, Choi M, Choi YM, Crook JM, de Sousa PA, Dvorak P, Freund C, Firpo M, Furue MK, Gokhale P, Ha HY, Han E, Haupt S, Healy L, Hei DJ, Hovatta O, Hunt C, Hwang SM, Inamdar MS, Isasi RM, Jaconi M, Jekerle V, Kamthorn P, Kibbey MC, Knezevic I, Knowles BB, Koo SK, Laabi Y, Leopoldo L, Liu P, Lomax GP, Loring JF, Ludwig TE, Montgomery K, Mummery C, Nagy A, Nakamura Y, Nakatsuji N, Oh S, Oh SK, Otonkoski T, Pera M, Peschanski M, Pranke P, Rajala KM, Rao M, Ruttachuk R, Reubinoff B, Ricco L, Rooke H, Sipp D, Stacey GN, Suemori H, Takahashi TA, Takada K, Talib S, Tannenbaum S, Yuan BZ, Zeng F, Zhou Q. (2015). Points to consider in the development of seed stocks of pluripotent stem cells for clinical applications: International Stem Cell Banking Initiative (ISCBI). Regen Med. 2015;10(2 Suppl):1-44.
48. Yusa K, L Zhou, MA Li, A Bradley and NL Craig. (2011). A hyperactive piggyBac transposase for mammalian applications. Proc Natl Acad Sci U S A 108:1531-6.

Figure Legends

Figure 1. *PiggyBac* targeting at the *ADRB2* locus. Panel (A) shows a schematic of the *ADRB2* locus structure before targeting, after insertion of the *PiggyBac* positive-negative selection cassette and after cassette excision. The black (G/G) and red (A/C) vertical lines indicate the location of the polymorphic changes induced at bases 46 and 79. Primer locations (b1, b2) for genotyping are indicated, along with PCR product sizes. β2-L and β2-R indicate the left and right regions of homology, each of 1kb in length; TV, targeting vector; PB, *PiggyBac*; PGK, phosphoglycerate kinase promoter; PURO, puromycin-N-acetyltransferase; TK, thymidine kinase. Panel (B) shows the time line of the conventional two step *PiggyBac* targeting approach (upper) and the simplified approach (lower). In (C), expression of the *ADRB2* gene was evaluated by quantitative realtime PCR in undifferentiated hPSCs (U) and through a 66 day timecourse of directed monolayer differentiation to cardiomyocytes; beating sheets appeared from between d8-12. Data are mean±SEM; n=4.

Figure 2. Polymorphic changes to the *ADRB2* locus in hPSCs. The nucleotide and translated single letter amino acid sequences are shown for the 5' region of the *ADRB2* locus. The targeting strategy introduces changes at positions 46 and 79 (non-synonymous in the peptide), and 124 and 126 (synonymous in the peptide) as indicated. Features identified are the location of the gRNA underlined, with PAM site boxed, and *TTAA PiggyBac* cassette insertion site.

Figure 3. Gene editing at the *ADRB2* locus in hPSCs. Panel (A) shows representative chromatogram synopses flanking positions 46, 79 and 124-126 of untargeted, mono- and bi-allelic targeting, and indels. A complete set for step 1 and step 2 targeting is in Supp Figs. 2A and 2B. The table in (B) summarises the different targeting events identified after step 1 (midpoint; after puromycin selection for clones containing the positive-negative selection cassette) and step 2 (after ganciclovir selection for clones in which the cassette has been excised). In panel (C), high risk off target (OT) sites were classified as known coding or regulatory sequences where gRNAs had full PAM site complementarity and/or fewer than 5 mismatches with the target. PCR genotyping showed no evidence for off target events. In (D), random integration was tested. The schematic shows the stages of targeting and location of PCRed regions. *ADRB2* is a control for genomic DNA, whilst *ADBR2-PT* and *TK* test for the presence of the targeting cassette; results are shown in the gel images. Since no product was identified for *AMP* within the pBlueScript backbone, qPCR was used and compared against a positive control (pos) comprising plasmid DNA diluted to the equivalent of single copy gene level in HUES7 parental DNA. Housekeeping gene was *18S*, n=3±SD; **** P<0.001, Dunnett's test.

Figure 4. Polymorphic changes to the *GRK5, ACTC1* and *RYR2* loci in hPSCs. The nucleotide and translated single letter amino acid sequences are shown for each locus. The targeting strategy introduces non-synonymous changes as indicated. In each case, the gRNA, with PAM site boxed, spans an endogenous *TTAA* cassette insertion site, which eliminates the need for changes to spare the targeting vector from Cas9-mediated cleavage.

Figure 5. Gene editing at the *GRK5, ACTC1* and *RYR2* loci in hPSCs. The schematics in panel (A) show the loci for each gene before and after editing, with damaging (black to red; *ACTC1*), protective (red to black; *GRK5*) or rescue (red to black; *RYR2*) polymorphisms introduced. L and R represent the left and right regions of homology, while primer locations for g, a and r are indicated (full details in Supp. Table 1). TV, targeting vector. In (B), semi-quantitative RT-PCRs were carried out for each gene in undifferentiated hPSCs (Un) and cardiomyocytes at day 30 of differentiation (CM). Bands were quantified by densitometry and normalised to β-actin (*ACTB*) as a house keeping gene. M, marker; n=2, errors are ± SD. Panel (C) shows representative chromatogram synopses flanking polymorphic positions for each gene, while editing efficiencies are displayed in the tables in (D). Note that for *RYR2*, it is not possible to tell whether the event was mono- or bi-allelic, hence the ?? symbols.

Figure 6. Retention of pluripotency characteristics in the edited hPSC lines. Panels (A-C) shows assessment of pluripotency characteristics in undifferentiated cells from each of the edited line. This included (A) immunostaining for the transcription factor, OCT4 (green; inset with DAPI [blue] counterstaining), (B) flow cytometry for TRA-1-81 (blue), SSEA4 (green) and SSEA1 (red), relative to unstained (purple), and (C) G-banding karyotyping of 30 metaphase spreads per line, with a representative karyogram shown for each. In panel (D), directed monolayer differentiation produced cardiomyocytes of >88% purity, as gauged by immunostaining for α-actinin (red) relative to total nuclei count (DAPI, blue). Scale bar is 100μm; n = 2-4, SD.

Figure 7. Functional effects of chronic isoprenaline on GRK5-L41 and –Q41 hPSC-CMs. Using the CardioExcyte impedance platform, the beat rate of the edited hPSC-CM lines was monitored at ~10 minute intervals during chronic stimulation (~50 h) with 100nM isoprenaline (Iso; Ai, Aii) with or without beta-blockade with 200nM propranolol (Prop; Bi, Bii). Data were binned for the periods shown and plotted as normalised to percent change from BL (baseline). The response ratios were calculated by dividing each datum from GRK5-Q41 by the corresponding time point from GRK5-L41 hPSC-CMs without (Aiii) or with (Biii) blockade with propranolol. Arrow head indicates where there is a highly significant decline in the beat

rate of the GRK5-Q41 hPSC-CMs. Dunnett's test relative to BL: * P<0.05; ** P<0.01; *** P<0.001; **** P<0.0001.

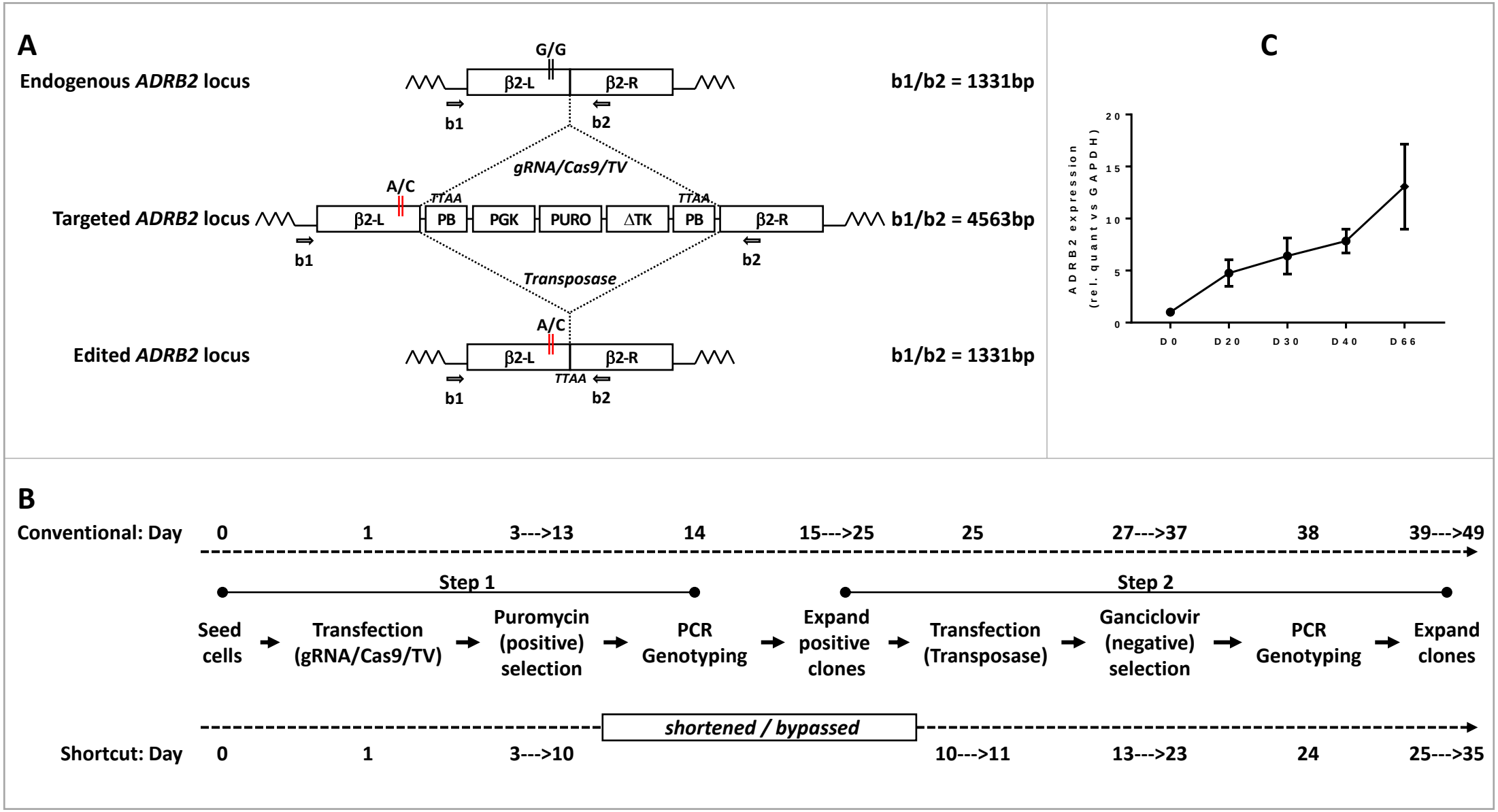

A
Endogenous ADRB2 locus
G/G
β2-L
β2-R
b1
b2
b1/b2 = 1331bp
gRNA/Cas9/TV
A/C
TTAA
TTAA
Targeted ADRB2 locus
β2-L
PB
PGK
PURO
ΔTK
PB
β2-R
b1/b2 = 4563bp
b1
b2
Transposase
A/C
Edited ADRB2 locus
β2-L
β2-R
TTAA
b1/b2 = 1331bp
b1
b2
C
ADRB2 expression (rel. quant vs GAPDH)
20
15
10
5
0
D0
D20
D30
D40
D66
B
Conventional: Day
0
1
3--->13
14
15--->25
25
27--->37
38
39--->49
Step 1
Step 2
Seed cells
Transfection (gRNA/Cas9/TV)
Puromycin (positive) selection
PCR Genotyping
Expand positive clones
Transfection (Transposase)
Ganciclovir (negative) selection
PCR Genotyping
Expand clones
shortened / bypassed
Shortcut: Day
0
1
3--->10
10--->11
13--->23
24
25--->35

Figure 1

ADRB2 (β2 adrenoceptor)

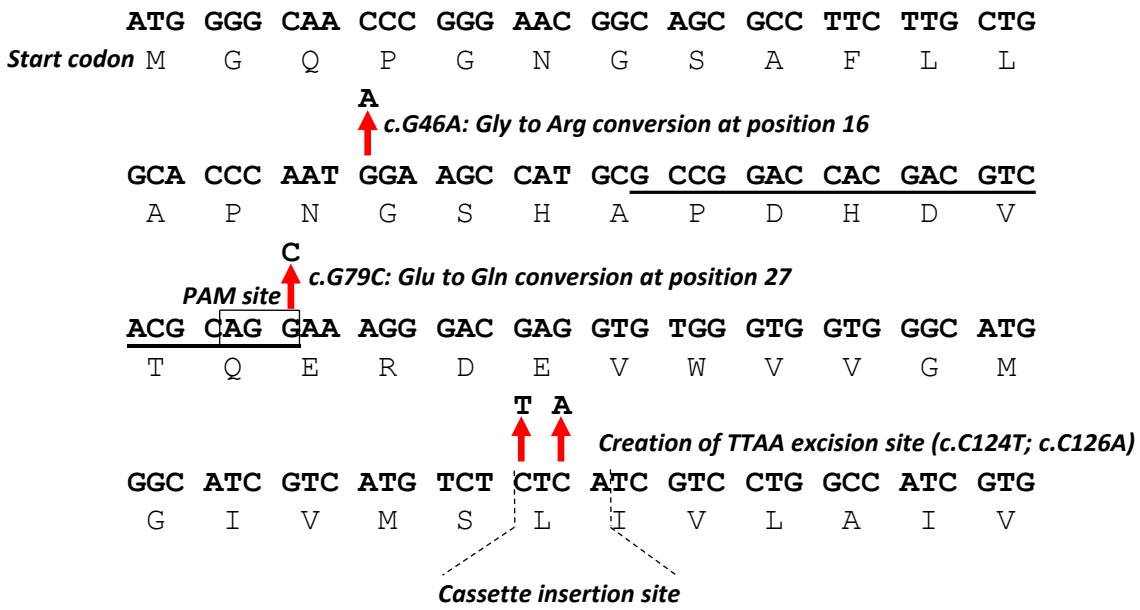
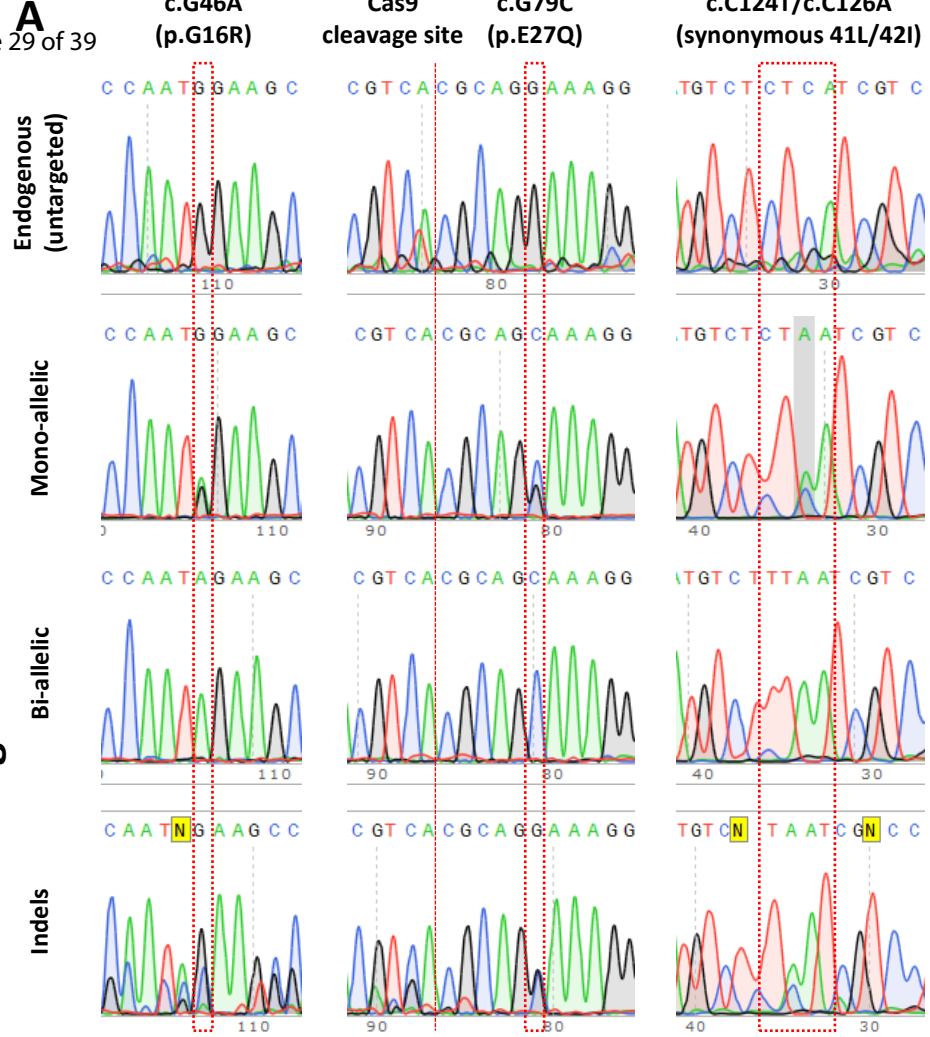


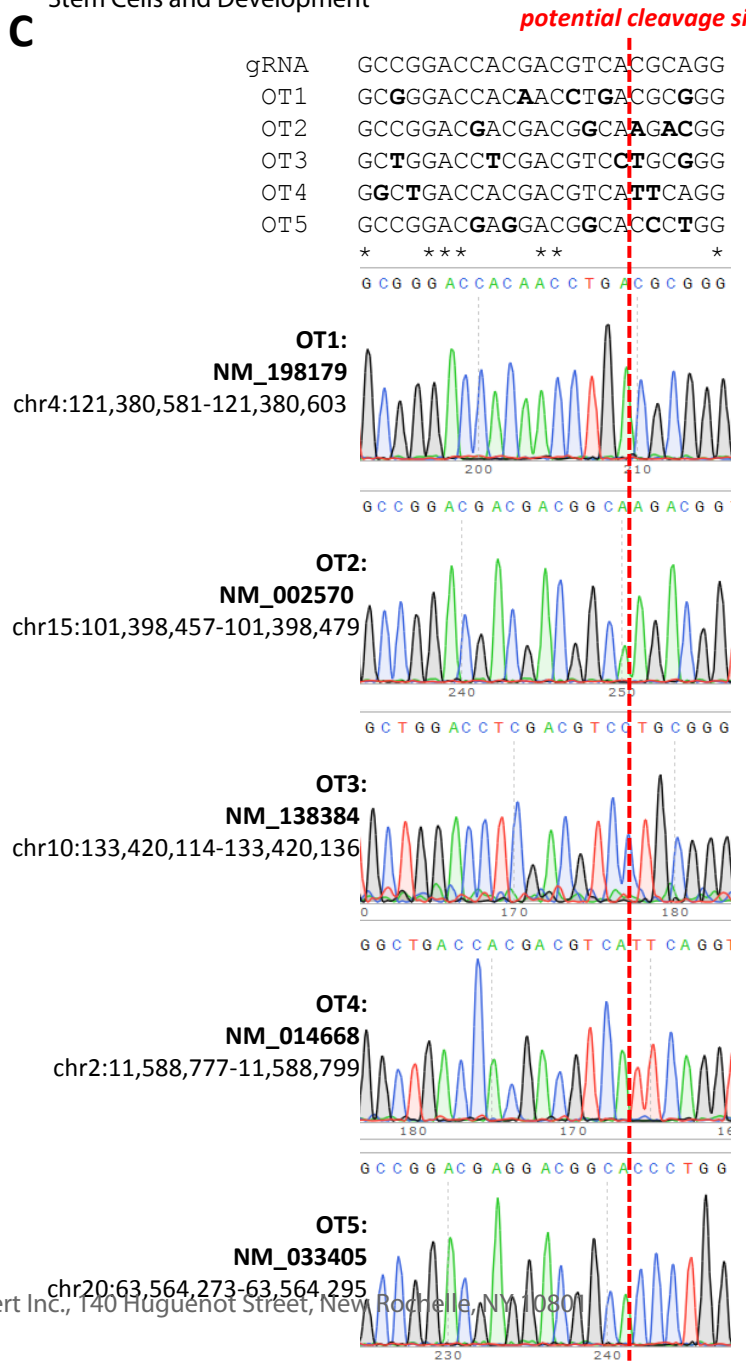
Figure 2

Stem Cells and Development



B

Event	1st step	2nd step
Untargeted	18.2% (2 of 11)	8.3% (1 of 12)
Mono-allelic	9.1% (1 of 11)	8.3% (1 of 12)
Bi-allelic	9.1% (1 of 11)	33.3% (4 of 12)
Indels	54.5% (6 of 11)	8.3% (1 of 12)
Unclear	9.1% (1 of 11)	41.7% (5 of 12)



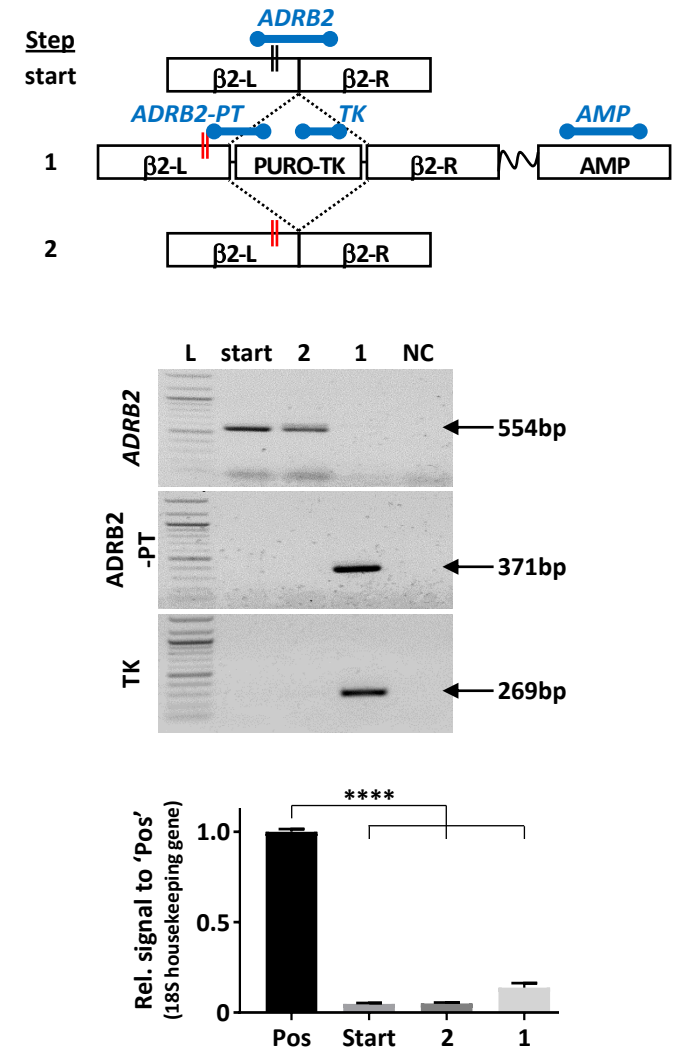


Figure 3

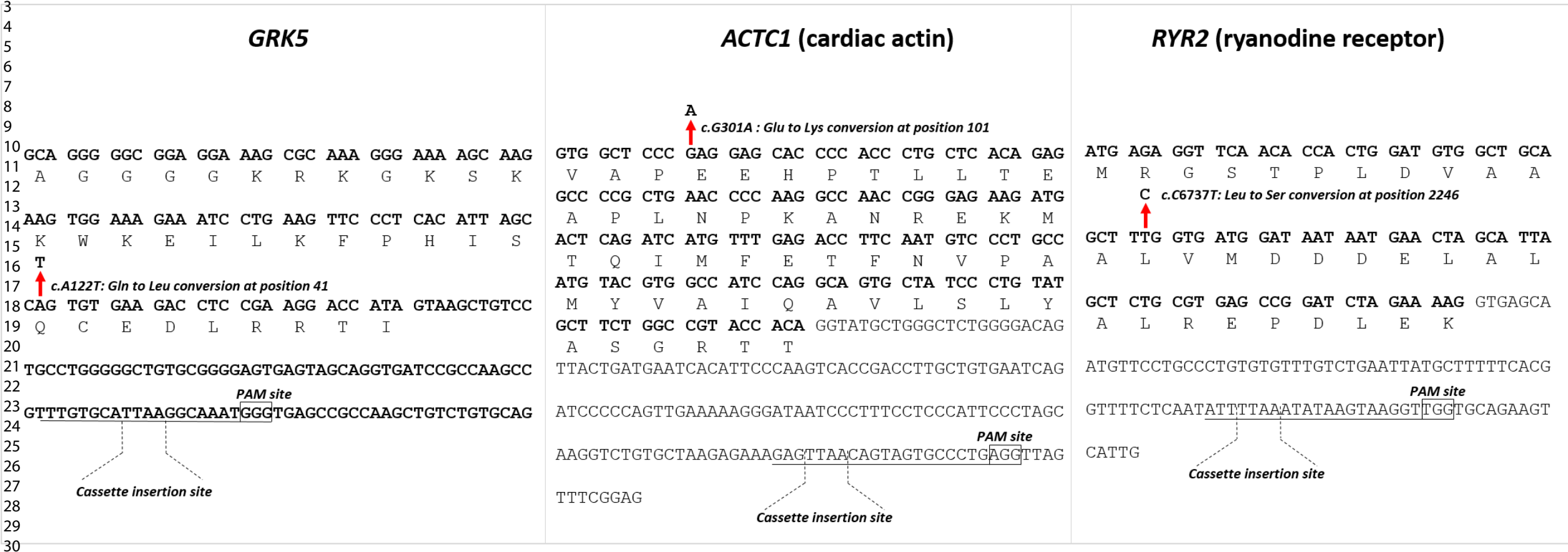

GRK5
GCA GGG GGC GGA GGA AAG CGC AAA GGG AAA AGC AAG
A G G G G K R K G K S K
AAG TGG AAA GAA ATC CTG AAG TTC CCT CAC ATT AGC
K W K E I L K F P H I S
T
c.A122T: Gln to Leu conversion at position 41
CAG TGT GAA GAC CTC CGA AGG ACC ATA GTAAGCTGTCC
Q C E D L R R T I
TGCCTGGGGGCTGTGCGGGGAGTGAGTAGCAGGTGATCCGCCAAGCC
PAM site
GTTTGTGCATTAAGGCAAATGGGTGAGCCGCCAAGCTGTCTGTGCAG
Cassette insertion site
ACTC1 (cardiac actin)
A
c.G301A : Glu to Lys conversion at position 101
GTG GCT CCC GAG GAG CAC CCC ACC CTG CTC ACA GAG
V A P E E H P T L L T E
GCC CCG CTG AAC CCC AAG GCC AAC CGG GAG AAG ATG
A P L N P K A N R E K M
ACT CAG ATC ATG TTT GAG ACC TTC AAT GTC CCT GCC
T Q I M F E T F N V P A
ATG TAC GTG GCC ATC CAG GCA GTG CTA TCC CTG TAT
M Y V A I Q A V L S L Y
GCT TCT GGC CGT ACC ACA GGTATGCTGGGCTCTGGGGACAG
A S G R T T
TTACTGATGAATCACATTCCCAAGTCACCGACCTTGCTGTGAATCAG
ATCCCCCAGTTGAAAAAGGGATAATCCCTTTCCTCCCATTCCCTAGC
PAM site
AAGGTCTGTGCTAAGAGAAAGAGTTAACAGTAGTGCCCTGAGGTTAG
TTTCGGAG
Cassette insertion site
RYR2 (ryanodine receptor)
ATG AGA GGT TCA ACA CCA CTG GAT GTG GCT GCA
M R G S T P L D V A A
C
c.C6737T: Leu to Ser conversion at position 2246
GCT TTG GTG ATG GAT AAT AAT GAA CTA GCA TTA
A L V M D D D E L A L
GCT CTG CGT GAG CCG GAT CTA GAA AAG GTGAGCA
A L R E P D L E K
ATGTTCCTGCCCTGTGTGTTTGTCTGAATTATGCTTTTTCACG
PAM site
GTTTTCTCAATATTTTAAATATAAGTAAGGTTGGTGCAGAAGT
CATTG
Cassette insertion site

Figure 4

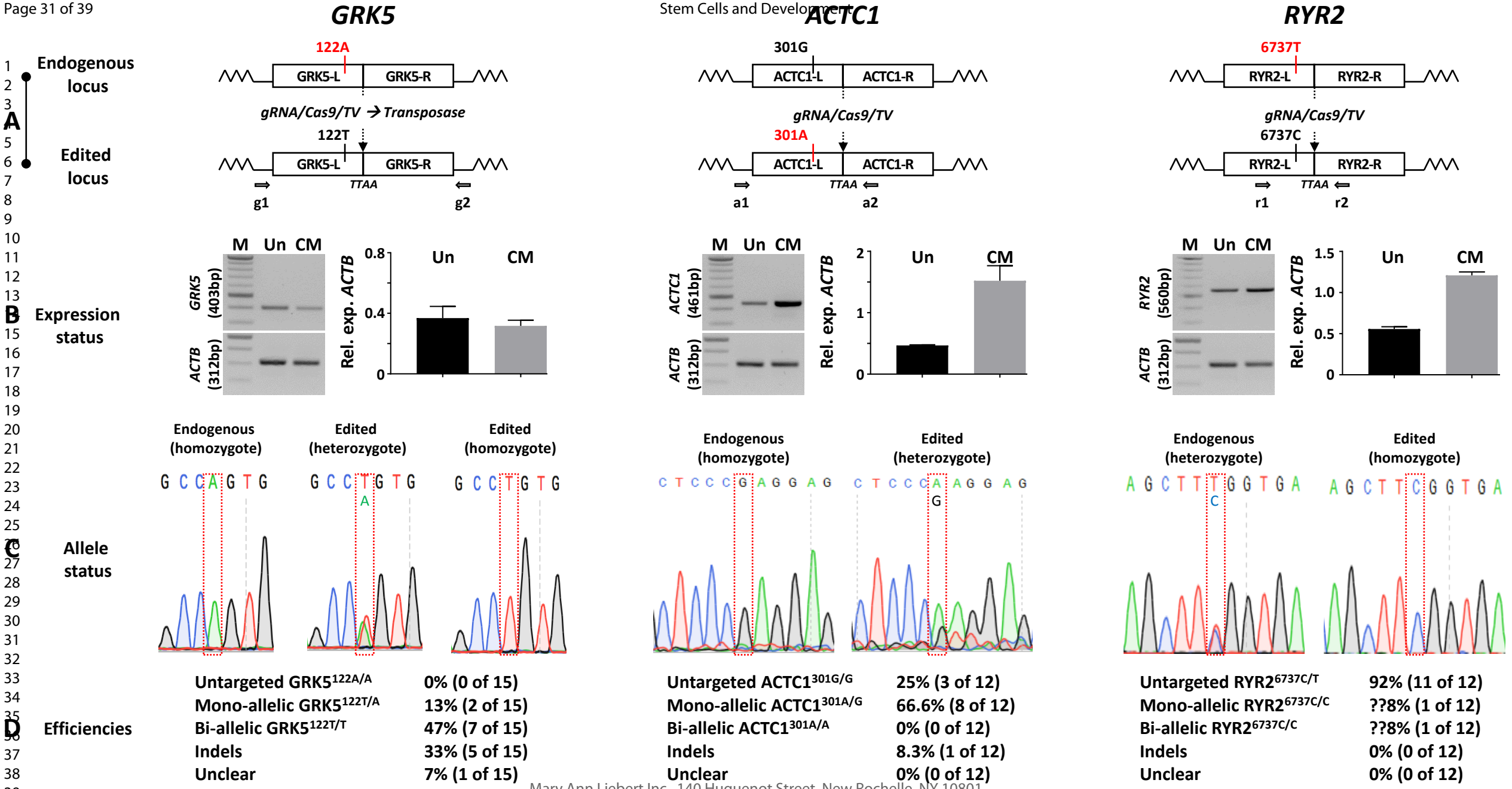

GRK5
ACTC1
RYR2
A
Endogenous locus
Edited locus
122A
GRK5-L
GRK5-R
gRNA/Cas9/TV → Transposase
122T
TTAA
g1
g2
301G
ACTC1-L
ACTC1-R
gRNA/Cas9/TV
301A
a1
a2
6737T
RYR2-L
RYR2-R
6737C
r1
r2
B
Expression status
M Un CM
GRK5 (403bp)
ACTB (312bp)
Rel. exp. ACTB
Un
CM
ACTC1 (461bp)
RYR2 (560bp)
C
Allele status
Endogenous (homozygote)
Edited (heterozygote)
Edited (homozygote)
Endogenous (heterozygote)
D
Efficiencies
Untargeted GRK5122A/A 0% (0 of 15)
Mono-allelic GRK5122T/A 13% (2 of 15)
Bi-allelic GRK5122T/T 47% (7 of 15)
Indels 33% (5 of 15)
Unclear 7% (1 of 15)
Untargeted ACTC1301G/G 25% (3 of 12)
Mono-allelic ACTC1301A/G 66.6% (8 of 12)
Bi-allelic ACTC1301A/A 0% (0 of 12)
Indels 8.3% (1 of 12)
Unclear 0% (0 of 12)
Untargeted RYR26737C/T 92% (11 of 12)
Mono-allelic RYR26737C/C ??8% (1 of 12)
Bi-allelic RYR26737C/C ??8% (1 of 12)
Indels 0% (0 of 12)
Unclear 0% (0 of 12)

Figure 5

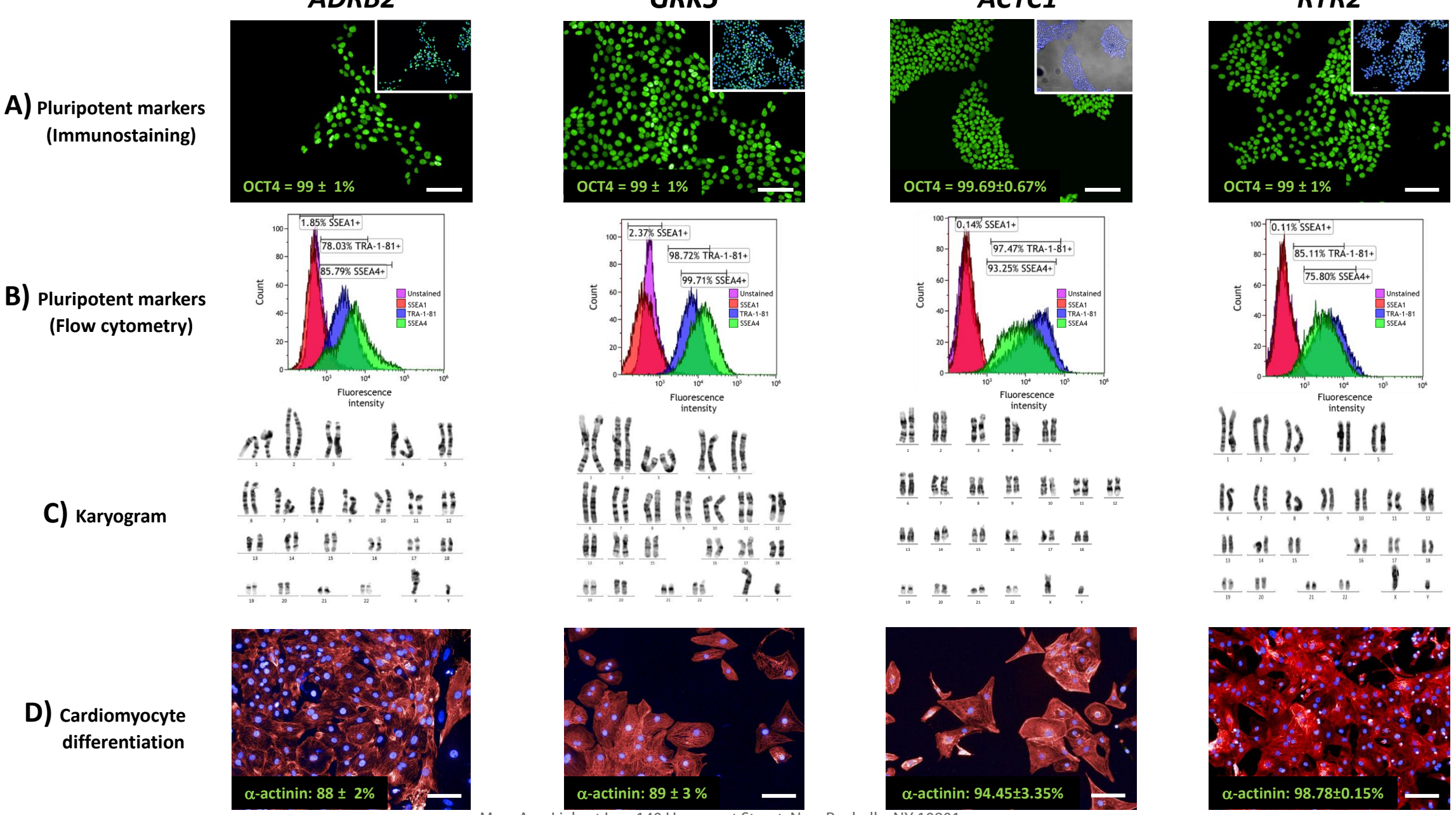

ADRB2
GRK5
ACTC1
RYR2
A) Pluripotent markers (Immunostaining)
OCT4 = 99 ± 1%
OCT4 = 99 ± 1%
OCT4 = 99.69±0.67%
OCT4 = 99 ± 1%
B) Pluripotent markers (Flow cytometry)
1.85% SSEA1+
78.03% TRA-1-81+
85.79% SSEA4+
2.37% SSEA1+
98.72% TRA-1-81+
99.71% SSEA4+
0.14% SSEA1+
97.47% TRA-1-81+
93.25% SSEA4+
0.11% SSEA1+
85.11% TRA-1-81+
75.80% SSEA4+
Unstained
SSEA1
TRA-1-81
SSEA4
Count
Fluorescence intensity
C) Karyogram
D) Cardiomyocyte differentiation
α-actinin: 88 ± 2%
α-actinin: 89 ± 3 %
α-actinin: 94.45±3.35%
α-actinin: 98.78±0.15%

Figure 6

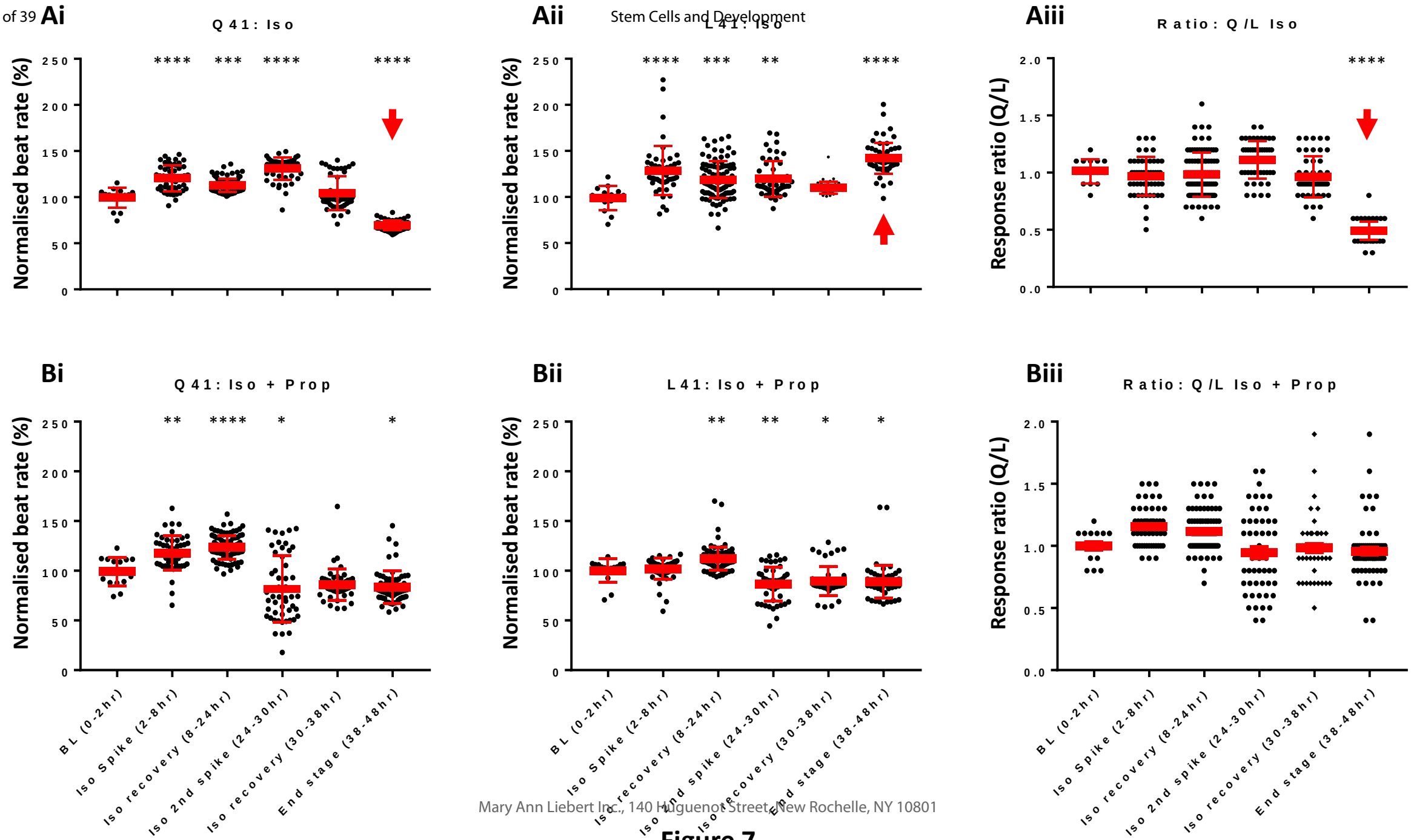

Ai
Q41: Iso
Normalised beat rate (%)
Aii
L41: Iso
Normalised beat rate (%)
Aiii
Ratio: Q/L Iso
Response ratio (Q/L)
Bi
Q41: Iso + Prop
Normalised beat rate (%)
Bii
L41: Iso + Prop
Normalised beat rate (%)
Biii
Ratio: Q/L Iso + Prop
Response ratio (Q/L)
BL (0-2hr)
Iso Spike (2-8hr)
Iso recovery (8-24hr)
Iso 2nd spike (24-30hr)
Iso recovery (30-38hr)
End stage (38-48hr)

Figure 7

Simplified footprint-free Cas9/CRISPR editing of cardiac-associated genes in human pluripotent stem cells

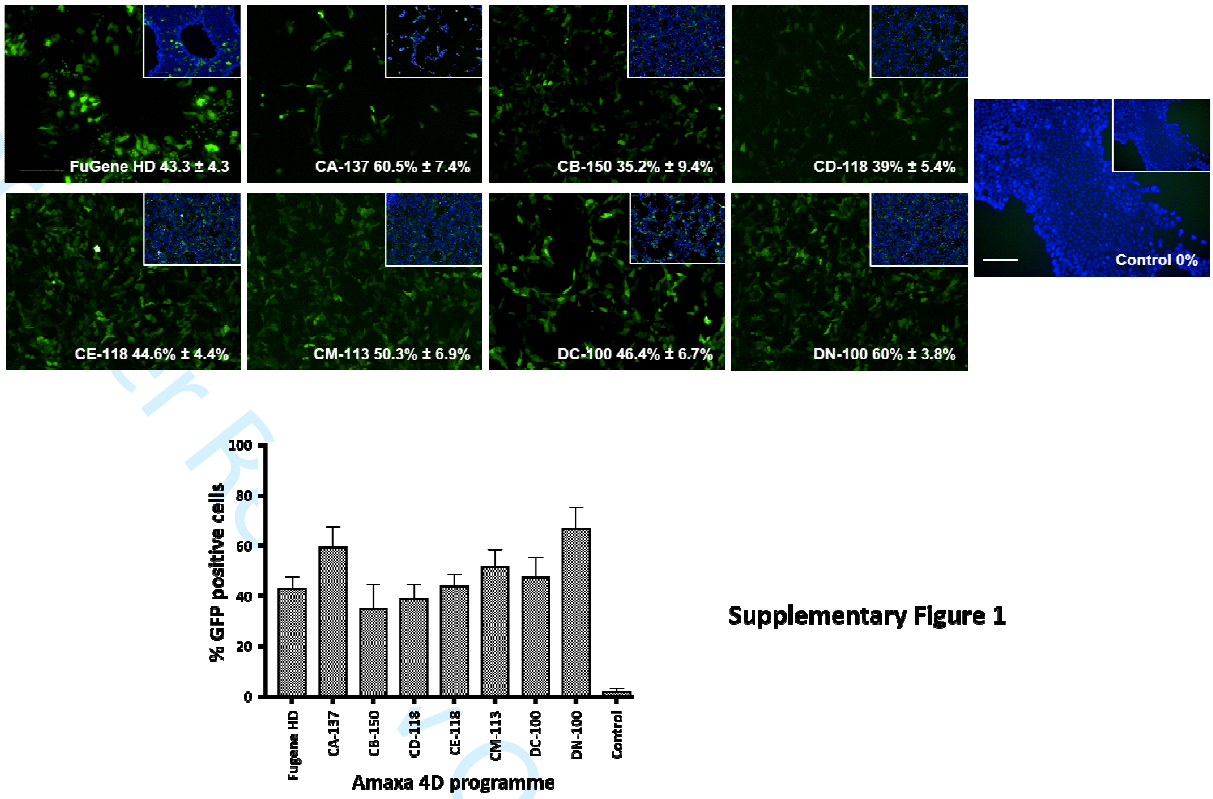
Alexander Kondrashov*†, Minh Duc Hoang†, James G.W. Smith, Jamie R. Bhagwan, Gary Duncan, Diogo Mosqueira, Maria Barbadillo Munoz, Nguyen T.N. Vo, Chris Denning*

Department of Stem Cell Biology, Centre of Biomolecular Sciences, University of Nottingham, NG7 2RD. United Kingdom

† Authors contributed equally to work

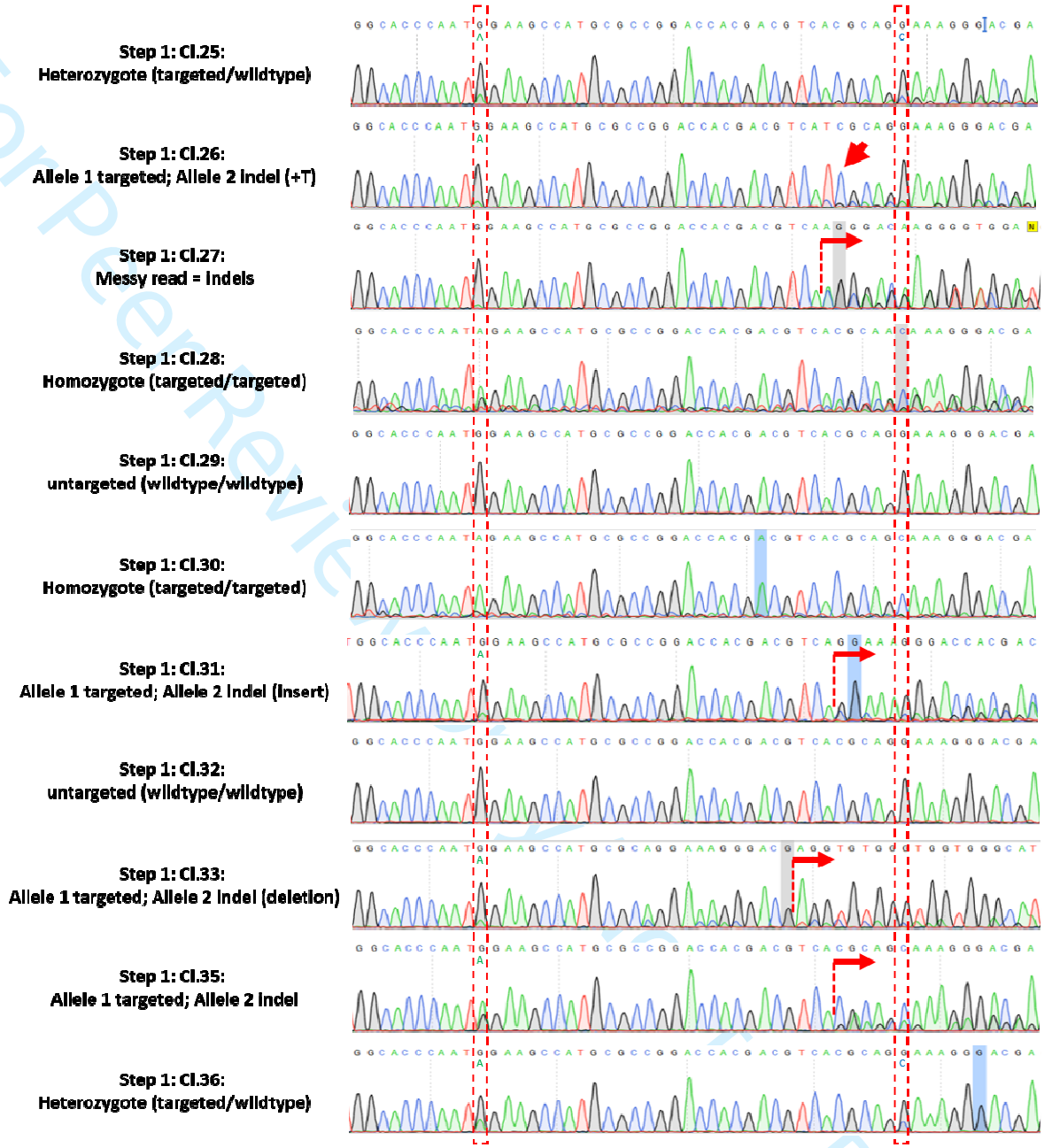
* To whom correspondence should be addressed.
Tel: +44(0)115 8231233; Fax: +44(0)115 8231230; Email: a.kondrashov@nottingham.ac.uk
Tel: +44(0)115 8231236; Fax: +44(0)115 8231230; Email: chris.denning@nottingham.ac.uk

Short title: Simplified CRISPR editing in hPSCs

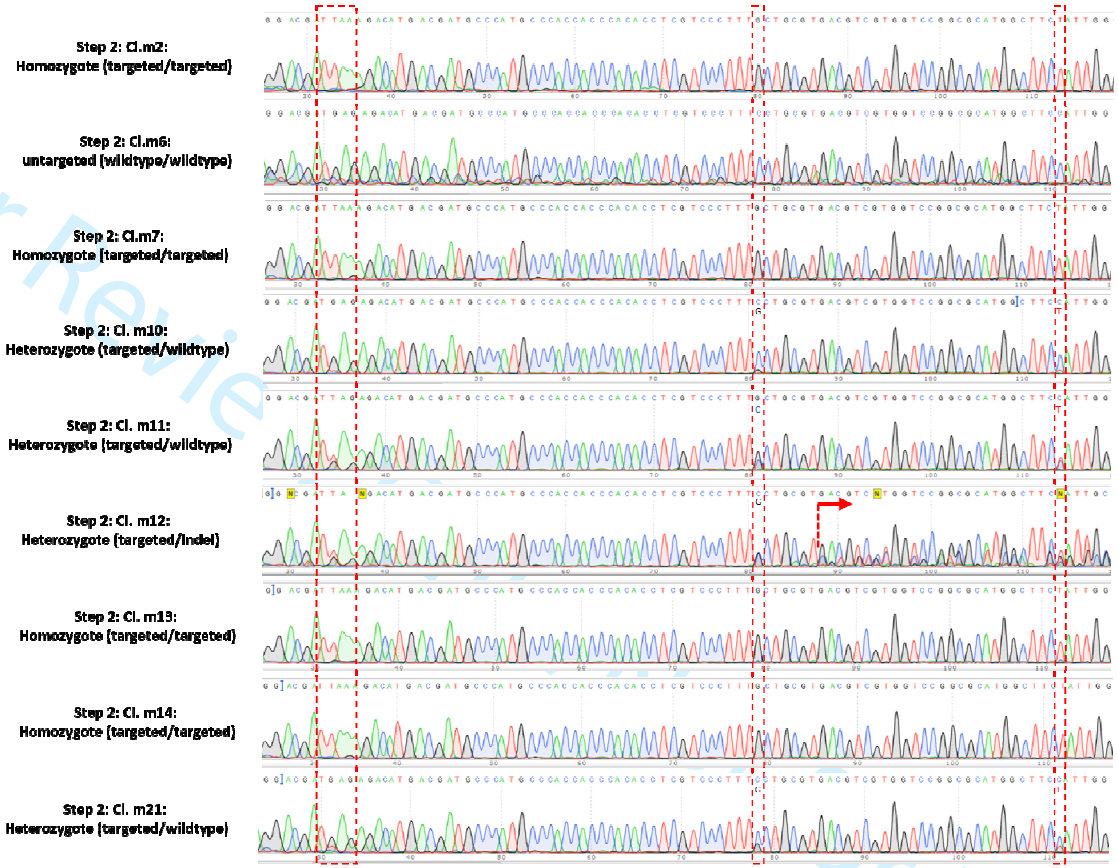


Supplementary Figure 1

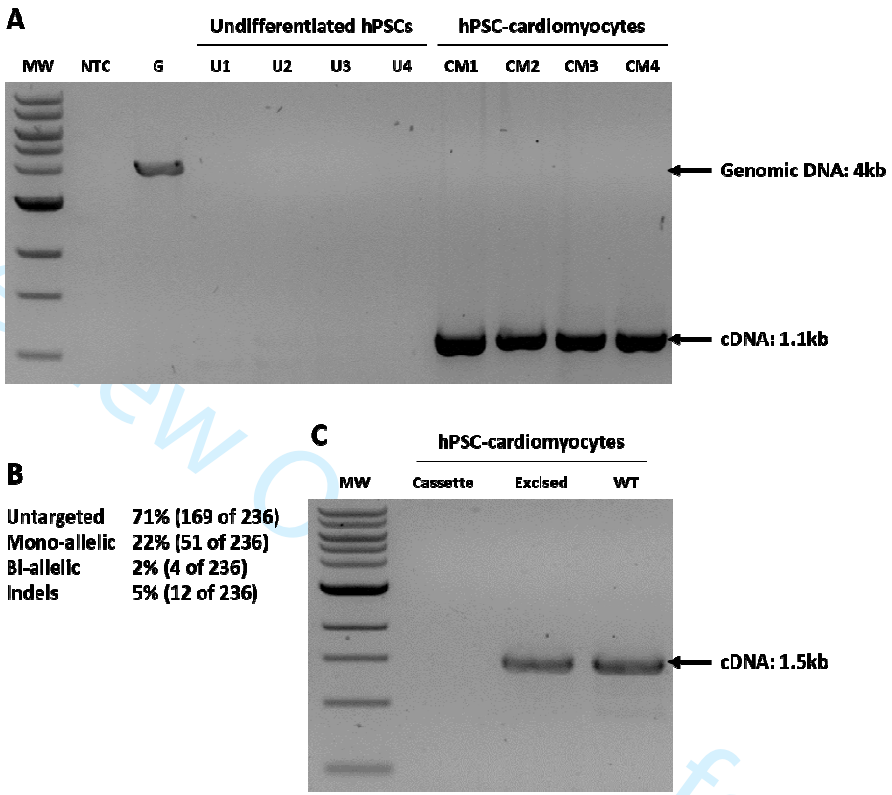
Supp Figure 1. Optimisation of transfection conditions for gene editing in hPSCs. Transfection efficiency of a GFP-expressing plasmid into undifferentiated HUES7 hESC was evaluated using chemical transfection with Fugene HD or using various programmes embedded in the Amaxa 4D nucleofection system. Shown are representative fields of GFP (green; inset is with DAPI [blue]), which were quantified by high content image analysis using the Operetta confocal plate reader. Data are n=3, mean ± SEM. Scale bar is 100µm.



Supp Figures 2A and 2B. Sequencing data for *ADRB2* editing. Chromatograms derived from direct sequencing of PCR products are shown for the clones picked after step 1 (Figure 2A; midpoint; after puromycin selection) and after step 2 (Figure 2B; end; after ganciclovir selection and cassette excision). In 2A (read in forward direction), the polymorphic sites at nucleotide positions 46 and 79 are shown, but not the *TTAA* site since this is disrupted by cassette insertion. In 2B (read in reverse direction), the polymorphic sites at nucleotide positions 46 and 79, as well as the reconstituted *TTAA* site. Heterozygote (mono-allelic) and homozygote (bi-allelic) targeting events are shown, with the former indicated by double peaks and both bases shown. Arrows indicate region where sequence becomes misaligned due to indels.



Supp Figures 2A and 2B. Sequencing data for *ADRB2* editing. Chromatograms derived from direct sequencing of PCR products are shown for the clones picked after step 1 (Figure 2A; midpoint; after puromycin selection) and after step 2 (Figure 2B; end; after ganciclovir selection and cassette excision). In 2A (read in forward direction), the polymorphic sites at nucleotide positions 46 and 79 are shown, but not the *TTAA* site since this is disrupted by cassette insertion. In 2B (read in reverse direction), the polymorphic sites at nucleotide positions 46 and 79, as well as the reconstituted *TTAA* site. Heterozygote (mono-allelic) and homozygote (bi-allelic) targeting events are shown, with the former indicated by double peaks and both bases shown. Arrows indicate region where sequence becomes misaligned due to indels.



Supplementary Figure 3

Supp Figure 3. Gene editing in the *MYH7* locus (β-myosin heavy chain). Panel (A) shows RT-PCR analysis of *MYH7* in different samples of hPSC lines in undifferentiated state (U1-U4) or after directed monolayer differentiation to cardiomyocytes (CM1-CM4); only the latter show expression. MW, molecular weight marker; NTC, no template control; G, genomic DNA. Despite the lack of gene expression in the undifferentiated hPSCs, panel (B) shows the targeting efficiencies of ~25% after insertion of a selection cassette flanked by FRT recombination sites. In (C), note that by RT-PCR *MYH7* expression occurs in wildtype (WT) cells or targeted cells from which the selection cassette has been removed with FLP recombinase (excised) but not when the selection cassette is present (cassette), even though it is positioned in an intron away from any annotated elements or splice junctions. MW, molecular weight marker.

Supp Table 1. Primers for vector construction and genotyping. LH, left homology; RH, right homology; VC, vector construction; CA, cassette amplification

Target	Accession	Reagent	Use	Forward / reverse	5′ → 3′ sequence	Product size (bp)	See Fig.
ADRB2	NM_000024	LH primer	VC	F	GGTCGACGGTATCGATAAGCTTGATTTCGGAGTACCCAGATGGAGAC	1084	Fig. 1A
				R	ACGCAGACTATCTTTCTAGGGTTAAAGACATGACGATGCCCATGC		
		RH primer		F	CAATATGATTATCTTTCTAGGGTTAATCGTCCTGGCCATCGTGTTTGG	1083	
				R	ATCCCCCGGGCTGCAGGAATTCGATAGTCTTCCGTGCCTGGGAGGTC		
		ADRB2puroΔtk	CA	F	CATCGTCATGTCTTTAACCCTAGAAAGATAGTCTGCG	3267	
		puroΔtkADRB2		R	ACACGATGGCCAGGACGATTAACCCTAGAAAGATAATCATATTGTGACG		
		Primer b1	Genotyping	F	GCTCGGGTGAGGCAAGTTCGG	1331	
		Primer b2		R	ATGGCAAAGTAGCGATCCAC		
		Primer b3	Expression	F	GCTGAGTGTGCAGGACGAGT	555	Fig. 5
		Primer b4		R	ATGGCAAAGTAGCGATCCAC		
		gRNA	Targeting	+ Strand	GCCGGACCACGACGTCACGC	n.a.	Fig. 2
	NM_198179	Primer OT1	OT1	F	CCTGAGCTGCTCTCCTTTCC	378	Fig. 3C
				R	CCAGAGCATTGCCAAAGAGC		
	NM_002570	Primer OT2	OT2	F	CAGCATGGAGAAGAGGAGCC	387	
				R	ACTGTCACCCTTGTCCCAGA		
	NM_138384	Primer OT3	OT3	F	GATGCCATCATGGAGCCTCT	357	
				R	ACCCTAGTGACCAGCATGGA		
	NM_014668	Primer OT4	OT4	F	TGCACTCAATGAGCAAGGCT	344	
				R	CCCAGCTGGACCAGGTAGTA		
	NM_033405	Primer OT5	OT5	F	GAGGAGGTGATCAGGCAGC	332	
				R	TCCCCATGCTTCTCACACAG		
	NM_000024	ADRB2	Random integration	F	GCTGAGTGTGCAGGACGAGT	554	Fig. 3D
				R	ATGGCAAAGTAGCGATCCAC		
	n/a	ADRB2-PT		F	GCTGAGTGTGCAGGACGAGT	371	
				R	CTAAATGCACAGCGACGGATTCGCGC		
		TK		F	ATAGACGGTCCTCACGGGAT	269	
				R	ATATGAGGAGCCAGAACGGC		
		AMP		F	CTGCAATGATACCGCGAGAC	576	
				R	TCCTTGAGAGTTTTCGCCCC		
GRK5	NM_005308	LH primer	VC	F	GGTCGACGGTATCGATAAGCTTGATTAAGCGCCACTGTAAGGGTGGAGAG	1156	Fig. 5
				R	ACGCAGACTATCTTTCTAGGGTTAATGCACAAACGGCTTGGCGGATCACC		
		RH primer		F	CAATATGATTATCTTTCTAGGGTTAAGGCAAATGGGTGAGCCGCCAAGCTG	1146	
				R	AGTGGATCCCCCGGGCTGCAGGAATTCGATGGCACAGATGGCCTCCTATC		
		Primer g1	Genotyping	F	CAGAGCAAGGTGGAGGACAG	1446	
		Primer g2		R	GATAGGAGGCCATCTGTGCC		
		Primer g3	Expression	F	TCCGAAGGACCATAGACAGAGA	403	
		Primer g4		R	TGCCTTTCCAACCACTTCCA		
	WGE ID: 1075072174	gRNA 2a	Targeting	- Strand	TTTGTGCATTAAGGCAAATG	n.a.	Fig. 4
	WGE ID: 1075072180	gRNA 2b		+ Strand	GTGAGGCAAATGCCAATCAG		
ACTC1	NM_005159	LH primer	VC	F	GGTCGACGGTATCGATAAGCTTGATGCCAGACAGGCTGCCAAGCAGG	1050	Fig. 5
				R	ACGCAGACTATCTTTCTAGGGTTAACTCTTTCTCTTAGCACAGAC		
		RH primer		F	CAATATGATTATCTTTCTAGGGTTAACAGTAGTGCCCTGAGGTTAGTTT	1051	

				R	ATCCCCCGGGCTGCAGGAATTCGATGCTGGAAGAGTGTCTCAGGACAG		
		Primer a1	Genotyping	F	CACCTGACCCTCTTGTTCGA	2267	
		Primer a2		R	GCGGATTCAGTGAGAGAGGA		
		Primer a3	Expression	F	GGTGATGAAGCCCAGAGCAA	461	
		Primer a4		R	GTGGTGACAAAGGAGTAGCC		
		gRNA	Targeting	+ Strand	GAGTTAACAGTAGTGCCCTG	n.a.	Fig. 4
RYR2	NG_008799.2	LH primer	VC	F	ACGGTATCGATAAGCTTGATTACGTAAAATTAAACTTTAA	360	Fig. 5
				R	GACTATCTTTCTAGGGTTAAAATATTGAGAAAACCGTGAA		
		RH primer		F	TGATTATCTTTCTAGGGTTAATATAAGTAAGGTTGGTGCA	1041	
				R	CCGGGCTGCAGGAATTCGATCAGTAAAGGAAACAGGAAGA		
		Primer r1	Genotyping	F	CCCCAGCTATGAGAGGTTCA	445	
		Primer r2		R	GAACGTTGGTTCTCCTTCCA		
		Primer r3	Expression	F	TGCATGAAAGCATCAAACGCA	560	
		Primer r4		R	TGAGTAGAGCCGGAGAGTGT		
		gRNA	Targeting	+ Strand	ATTTTAAATATAAGTAAGGT	n.a.	Fig. 4
ACTB	NM_001101.3	Primer GAP1	Expression	F	CAAGAGATGGCCACGGCT	312	Fig. 5
		Primer GAP2		R	CTTGATCTTCATTGTGCTGGG		