Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

Article

Using High-Content Screening Technology as a Tool to Generate Single-Cell Patient-Derived Gene-Corrected Isogenic iPS Clones for Parkinson's Disease Research

Peter A. Barbuti [1,2,3,*], Paul M. Antony [2], Bruno FR. Santos [1,2], Francois Massart [2], Gérald Cruciani[2], Claire M. Dording [1,2], Jonathan Arias [4,5], Jens C. Schwamborn [4], Rejko Krüger [1,2,6,7,*,§]

1 Transversal Translational Medicine, Luxembourg Institute of Health, L-1445, Luxembourg; peter.barbuti@lih.lu (P.B.); bruno.santos@uni.lu (B.S.); claire.dording@lih.lu (C.D.); rejko.krüger@lih.lu (R.K.).
2 Clinical and Experiment Neuroscience, Luxembourg Centre for Systems Biomedicine, University of Luxembourg, L-4362 Luxembourg; paul.antony@uni.lu (P.A.); francois.massart@uni.lu (F.M.); gerald.cruciani@uni.lu (G.C.)
3 Department of Pathology and Cell Biology, Columbia University Irving Medical Center, 10032, New York, NY, USA
4 Developmental and Cellular Biology, Luxembourg Centre for Systems Biomedicine, University of Luxembourg, L-4362 Luxembourg; jonathan.arias@uni.lu (J.A.); jens.schwamborn@uni.lu (J.C.)
5 Department of Biosciences and Nutrition, Neo, Karolinska Institutet, Huddinge, Sweden.
6 Department of Neurodegenerative diseases, Hertie Institute for Clinical Brain Research, University Clinics Tübingen, Tübingen, Germany
7 Parkinson Research Clinic, Centre Hospitalier de Luxembourg (CHL), Luxembourg
* Correspondence: peter.barbuti@lih.lu (P.B); rejko.krüger@lih.lu (R.K.); Tel.: (+352-26970-967)
§ Senior author

Abstract: The generation of isogenic induced pluripotent stem cell (iPSC) lines using CRISPR-Cas9 technology is a technically challenging, time-consuming process with variable efficiency. Here we use fluorescence-activated cell sorting (FACS) to sort biallelic CRISPR-Cas9 edited single-cell iPS clones into high-throughput 96-well microtiter plates. We used high-content screening (HCS) technology and generated an in-house developed algorithm to select the correctly edited isogenic clones for continued expansion and validation. In our model we have gene-corrected the iPSCs of a Parkinson's disease (PD) patient carrying the autosomal dominantly inherited heterozygous c.88G>C mutation in the *SNCA* gene, which leads to the pathogenic p.A30P form of the alpha-synuclein protein. Undertaking a PCR restriction-digest mediated clonal selection strategy prior to sequencing, we were able to post-sort validate each isogenic clone using a quadruple screening strategy. Subsequent transfection with mRNA encoding excision-only transposase allows for the generation of footprint-free isogenic iPSC lines. These monoclonal isogenic iPSC lines retain a normal molecular genotype, express pluripotency markers and have the ability to differentiate into the three germ layers. This combinatory approach of FACS, HCS and post-sorted restriction digestion facilitates the generation of isogenic cell lines for disease modelling to be scaled-up on an automated platform.

Keywords: CRISPR-Cas9; high-content screening (HCS); Fluorescent-activated cell sorting (FACS); Parkinson's disease (PD), patient-derived iPS; single-cell clones; isogenic cell lines; *SNCA*; alpha-synuclein; A30P;

© 2020 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

1. Introduction

Parkinson's disease (PD) is the most common neurodegenerative movement disorder. The global burden of PD, assessed in 2016 was 6.1 million patients, this is estimated to reach 12 million patients worldwide by 2050 [1]. PD is clinically characterised by the two neuropathological features, the degeneration of innervating A9 dopaminergic neurons from the substantia nigra pars compacta (SNc) to the striatum in the midbrain, and the formation of intra-cytoplasmic neuronal inclusion bodies, referred to as Lewy Body's that are immunopositive for the alpha-synuclein protein in the neurons that remain [2].

PD has a heritability of between 10-15% with mutations in several genes explaining between 5-10% of these familial cases. The first PD gene identified was *SNCA* that encodes alpha-synuclein identified as an autosomal dominant form of the disease. Rare and highly penetrant missense mutations in the *SNCA* protein at p.A53T [3], p.A30P [4], p.E46K [5], p.G51D [6], and p.A53E [7] have all been identified, together with duplications [8] and triplications [9] of the *SNCA* gene locus. For the latter higher protein levels of alpha-synuclein correlated to increased severity and progression of the disease. Moreover, different Genome Wide Association Studies (GWAS) have identified common variations in *SNCA* as susceptibility factors for PD in populations worldwide [10,11].

The advancement in gene-editing by utilising the RNA guided Cas9 nuclease from the clustered regularly interspaced short palindromic repeats (CRISPR) of the bacterial adaptive immune system in human cells has revolutionised disease modelling [12,13]. Genome editing by CRISPR-Cas9 generates a double-strand break (DSB) in which the error-free homology-directed repair (HDR) as opposed to the error-prone non-homologous end-joining (NHEJ) is used to repair the DNA. Consequently, this allows the gene editing and correction of pathological missense mutations to take place *in-vitro*, thereby isolate, and determine the exact effect of the specific mutation in relation to its isogenic corrected control.

The generation of gene-corrected patient-derived isogenic iPS cell lines in recent years typically involves groups successfully using antibiotic resistance combined with fluorescence activated cell sorting (FACS) technology, before embarking on a screening and sequencing campaign to select, if successful the isogenic clone [14,15]. Recently, the biallelic genomic editing technique has successfully used FACS in addition to antibiotic resistance to both introduce and gene-correct heterozygous mutations [16–18]. Although CRISPR-Cas9 gene knockout of single-cell HEK293T cells have been generated [19], and multiple clones of a gene corrected isogenic cell has been generated [20], to date there is no reported study of single-cell gene-corrected patient-derived isogenic iPSC clones.

Due to the technical complexity of gene-editing cultured human iPSCs, isogenic cell lines are generated as a polyclonal cell population, an advantage here is that higher cell numbers are achieved earlier in the derivation process. In contrast, a disadvantage of polyclonal isogenic cell lines is that the cells within the colony can have different proliferation rates. This is of particular importance where the gene editing concerns a developmental, cell-cycle or cell-death affected mutation, which over the course of the culture and repeated passaging, changing cellular composition can take place leading to variance to the research findings. Characterized single-cell clones offer the certainty of having a healthy genetic background, absent of biases in different proliferation rates. The quality control provided by the generation of single-cell isogenics ensures the reliability of phenotyping assays necessary for future drug discovery and translational research.

In this study, we gene-corrected the PD patient-derived iPS cells containing the heterozygous c.88G>C mutation in *SNCA* that generates the pathogenic A30P alpha-synuclein protein. We apply the biallelic genomic editing technique of antibiotic resistance selection followed by triple-reporter FACS technology to sort the CRISPR-Cas9 gene-edited iPSCs with a single-cell iPS clone per well of a 96-well plate. We then developed an algorithm and used OPERA® (Perkin Elmer) high-content screening (HCS) technology to post-sort screen the correctly edited clones prior to cell culture expansion. For clonal selection, we exploited that the c.88G>C p.A30P *SNCA* mutation itself leads to the creation of an *Mva1* restriction digest site within the exon 2 of the *SNCA* gene [4], thereby negating the need to generate additional silent mutations that would be required for the PCR-mediated clonal selection process. Thereby using PCR amplification and restriction-digestion, we use the undigested

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

PCR product as a way to validate the gene-corrected clones. Using this validation process of HCS and restriction digestion, we are able to quickly and cost-effectively generate and screen our single-cell iPSC clones (12 of the picked and amplified 34 clones were correctly edited, efficiency rate = 35.30%). We then performed Sanger sequencing as a validation step to confirm the gene editing, before excising the transcript and genotyping the footprint-free isogenic iPSC lines. Furthermore, we characterised these cells to ensure that the cells pluripotency ability is retained. We show that the generation of single-cell isogenic human iPSC lines assisted by the quadruple selection strategy (QSS) validation of: antibiotic resistance, single-cell FACS, PCR amplification and restriction-digestion mediated clonal selection, and finally Sanger sequencing is a process that facilitates the real-time tracking of the edited clones. This process not only increases the traceability of the *de-novo* single-cell isogenic cell line but increases the efficiency compared to previous methods [14,21]. Furthermore allowing the operator to track the successfully edited clone(s) through each stage of the QSS process aids accuracy, speed and efficiency of the isogenic iPSC generation, a current limitation within the genome-editing field. Importantly this strategy can be applied across multiple research areas within disease modelling.

2. Materials and Methods

Cell line and ethical approval

Skin biopsies were obtained after informed consent from a patient with Parkinson's disease. The patient carries a heterozygous mutation c.88G>C in the *SNCA* gene generating the pathogenic p.A30P form of the alpha-synuclein protein. This patient is an affected sibling of the index patient from the 1998 study [4,22]. The generation and characterisation of induced pluripotent stem cells (iPSCs) from the dermal fibroblasts has been described [23] and has a unique identifier HIHDNDi001-B (https://hpscreg.eu/cell-line/HIHDNDi001-B). Ethical approval for the development of and research pertaining to patient-derived cell lines have been given by the National Committee for Ethics in Research, Luxembourg (Comité National d'Ethique dans la Recherche; CNER #201411/05).

Cell culture

iPSCs were routinely cultured in 6-well plates (Nunc, 140675). These were coated with high concentration growth factor reduced Matrigel® (1:100; Corning, 354263) according to the manufacturer instructions. The iPSCs were maintained in homemade E8: (DMEM/F12 + HEPES)(Life Technologies; 31330038), Insulin-Transferrin-Selenium (ITS) (1%; Life Technologies, 41400045), Penicillin-Streptomycin (1%; Life Technologies, 15140), L-Ascorbic acid 2-phosphate sesquimagnesium salt hydrate ((AA2PM); 64µg/mL; Sigma, A8960), bFGF (10ng/mL; Peprotech, 100-18B), TGF-β1 (2ng/mL; Peprotech, 100-21), Heparin (100ng/mL; Sigma, H3149). The iPSCs were maintained as colonies and passaged in Dispase® (5U/mL; CellSystems, LS02104).

Generation of isogenic cell lines – insertion of fluorescent constructs

To generate isogenic stem cells, two biallelic constructs: the dTOMATO-T2A-Puromycin cassette (Addgene, 100604), and the EGFP-TA-Puromycin cassette (Addgene, 100603) containing the wild-type sequence is shown in Figure 1 in addition to the sgRNA 630 previously published [16]. Briefly, the iPSCs were dissociated to single cells using Accutase and plated in iPS media plus Rho-Kinase Inhibitor Y-27632 (10µM; Abcam, Ab120129). Then, $1x10^6$ iPSCs were electroporated once using the 2D-Amaxa nucleofector unit (Lonza) and the program B16. 1.5µg of each donor cassette and 2.5µg of sgRNA was used together for the electroporation. The Human Stem Cell Nucleofector Kit 1 was used for the transfection according to the manufacturer's instructions. After electroporation, 1mL of homemade E8 was added to the cuvette before being placed in the incubator for 10 minutes. The cells were then plated into three wells of a 6-well plate with approximately 1.5mL of media per well. Cell selection using antibiotic resistance to puromycin (Sigma, P9620) was used as soon as small to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

medium-sized colonies began to appear. Puromycin concentrations of 0µg/mL, 0.5µg/mL and 1µg/mL were used respectively in the three plated wells for 24 hours with the viable colonies following 1µg/mL antibiotic treatment passaged and then expanded into 2x 10cm² tissue culture treated Matrigel™-coated dishes (Nunc, Z755923) prior to sorting.

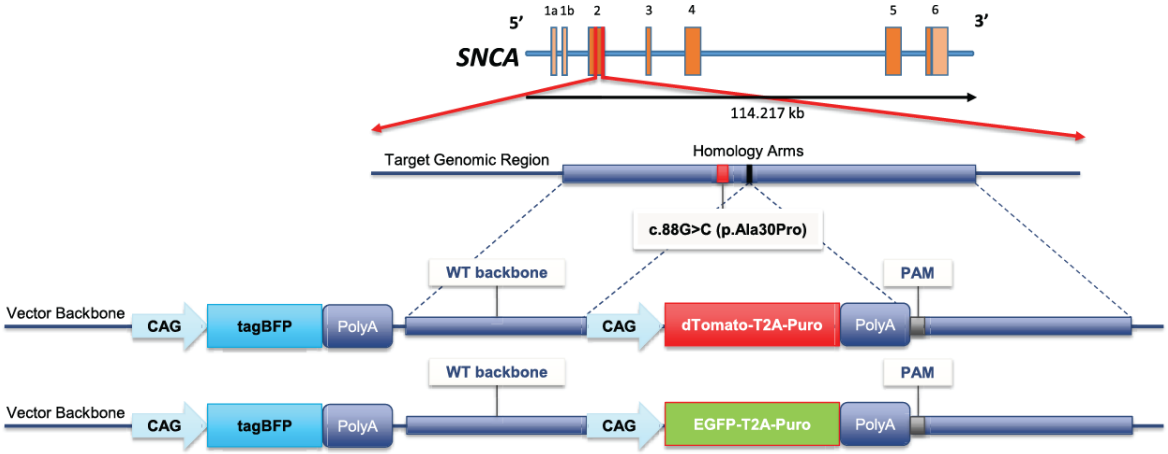

Figure 1. Donor Vector for Homology Directed Repair. The c.88G>C mutation is located in the target genomic region of Exon 2 of the *SNCA* gene located on chromosome 4. The vector backbone of the two constructs contain the tagBFP outside the homology arms, within the homology arms is the wildtype genomic DNA with the dTOMATO or EGFP fluorescent constructs.

Generation of isogenic cell lines – Sorting single-cell iPSCs

iPSC colonies were dissociated to single cells using Accutase before being centrifuged (300g; 3mins) and re-suspended in sterile-filtered sorting buffer (PBS containing EDTA (1mM; Sigma, E9884), HEPES (20mM, Life Technologies, 15630), Bovine Serum Albumin (0.2%, Sigma, A2058), 1% Penicillin/Streptomycin) and 10uM Rho-Kinase Inhibitor. Cells were sorted at 4˚C using an 85µM nozzle and a neutral density filter of 1.5 on the BD FACS Aria™ III (BD Biosciences). Cytometer Setup and Tracking (CST) beads were used daily to calibrate and define the baseline performance of the machine using the FACSDiva™ software. Prior to sorting, the drop delay experiment was setup to calibrate the sorting efficiency using the Accudrop™ beads, this was manually adjusted until the efficiency was ≥99.5%. In preparation of the single-cell sorting in a 96-well plate (Nunc, delta surface treated #167008) a dummy plate was placed onto the cooled stage to setup the sorting parameters. Briefly, the single-cell mode was selected in the FACSDiva™ software and 100 beads were sorted per well in order to visually check if the cells were sorted to the centre of the well, or if further calibration was required. Once the visual inspection was passed, a 96-well plate pre-coated with Matrigel, containing 100µL of Homemade E8 iPSC media and Rho-Kinase Inhibitor was placed onto the stage with 1 cell/well selected in the experimental setup parameters. The cells once sorted were left in the incubator (37°C, 5% CO_2) for 48 hours before the media being half-changed daily; Rho-Kinase Inhibitor was not included in the media at this stage. The cells were maintained for 10-21 days before they appeared visible and were able to be screened on the OPERA® High-Content Screening (HCS) System (Perkin Elmer).

Generation of isogenic cell lines – Expansion of the single-cell clones

Visual inspection was used to select the clones for subsequent passaging and expansion. These selected clones had an undifferentiated morphology, which was later confirmed by immunofluorescence, and had taken up both the dTOMATO and EGFP constructs; any clone that expressed the tagBFP construct was not expanded as a monoclonal iPS cell line. The selected clones were dissociated using Accutase and were expanded into a well of a 24-well plate before being

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

passaged into two wells of a 12-well plate. At this stage one of the cell clones was cryopreserved (70% Knockout Serum Replacement (KOSR), Life Technologies, 1867715; 20% Homemade E8, 10% DMSO, Sigma, D2438) with 10μM Rho-Kinase Inhibitor. The remaining clone was pelleted for PCR-mediated clonal selection using *Mva1* restriction digestion.

PCR Restriction digest mediated clonal selection

Total RNA was extracted from the iPSCs using NucleoSpin® RNA (Macherey-Nagel). RNA was treated with DNase I (Sigma, 047167280001) cDNA was reverse transcribed using Transcriptor High Fidelity cDNA Syntheisis Kit (Roche). Amplification of Exon 2 *SNCA* cDNA by PCR was performed using forward primer: CCCCGAAAGTTCTCATTCAA and reverse primer: TCCACCTTTTTGACAAGCAA giving a 314 bp product. Kit Go Taq® G2 Flexi DNA Polymerase (Promega M7805) was used for this reaction with the following program: Pre-denaturation (95˚C; 2mins), 35 cycles of denaturation (95˚C; 30s), annealing (60˚C; 30s) and extension (72˚C; 60s), followed by a final extension (72˚C; 5mins). *Mva1 (BstNI)* (30min, 37˚C; (ThermoFisher FastDigest #FD0554)) digestion was used to check clonal efficiency. The c.88G>C, p.A30P *SNCA* mutation leads to the creation of a *Mva1* restriction site [4], correction of this mutation results in the undigested product. Following the restriction digest mediated clonal selection, Sanger sequencing was performed on selected cell lines by Eurofins Genomics Germany GmBH.

Transposase-mediated generation of footprint-free isogenic cell lines

Both the dTOMATO-T2A-Puromycin cassette and the EGFP-TA-Puromycin cassette contain a TTAA sequence necessary for removal of the constructs [16]. The single cell clones containing the tagBFP$^-$/EGFP$^+$/dTOMATO$^+$ combination were transfected with an mRNA encoding excision-only transposase as previously described [16]. The Stemfect™ RNA Transfection Kit (Stemgent) was used according to the manufacturing instructions, this lead to the generation of footprint-free isogenic cell lines. Three sorting steps were undertaken to purify and confirm removal of the biallelic constructs (Supplementary Figure 3).

Immunocytochemistry

iPSCs were fixed, stained and imaged as previously described [24]. Primary antibodies used were: OCT4 (1;200; Santa Cruz, sc-5279), NANOG (1:100; Abcam, ab21624), SOX2(Y-17) (1:200; Santa Cruz, sc-17320), TRA-1-60 (1:300; Abcam, ab16288). Secondary antibodies used were: Alexa Fluor 488 Goat anti-Mouse IgG (H+L) (1:200; Invitrogen, A11029), Alexa Fluor 568 Goat anti-Rabbit IgG (H+L) (1:200; Invitrogen, A11036) and Alexa Fluo 647 Donkey anti-Goat IgG (H+L)(1:200; Invitrogen, A21447) .

***In-vitro* differentiation**

iPSCs were plated onto Geltrex™-coated glass coverslips. Directed *in-vitro* differentiation to the three germ layers was performed using the Human Pluripotent Stem Cell Functional Identification Kit (R&D Systems, SC027B) according to manufacturer instructions. Primary antibodies used provided with the kit were: SOX17 (1:1000; R&D Systems, #963121), OTX2 (1:1000; R&D Systems, #963273), BRACHYURY (1:1000; R&D Systems, #963427). The secondary antibody used was Alexa Fluor 647 Donkey anti-Goat IgG (H+L) (1:1000; Invitrogen, A21447). Cells were fixed, stained and imaged previously described [24].

Chromosomal analysis

Molecular karyotyping and identity analysis was performed on the iPS clones at Life&Brain GmBH (Bonn, Germany) using the Illumina BeadArray HumanOmni2.5Exome-8 BeadChip v1.3 on the Illumina iScan (Serial Number: N263) scanner (Illumina Inc. San Diego, USA). Genotype analysis

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

was performed using GenomeStudio V2.0.2 with Copy Number Analysis undertaken using the CNV-Partition V3.2 (Illumina Inc. San Diego, USA). Copy number events were reported if larger than 350,000 base pairs. The method overviewing the high-resolution whole sequence genomic profiling technology that was used in this study has been previously described [25]. The molecular karyotyping report for each iPS clone analysed are available with the authors and can be made available upon reasonable request.

Computer Code, software and licensing

Fluorescence microscopy images were acquired on an Opera QEHS spinning disc microscope (Perkin Elmer) using a 10x air objective with numerical aperture 0.4. Blue, green, and red fluorescent channels were acquired simultaneously. The sample was excited with 405 nm, 488 nm, and 561 nm lasers. Blue emission was detected behind a 450/50 bandpass filter, green emission behind a 520/35 filter, and red emission behind a 600/40 filter. The camera binning for all channels was 2.

Image analysis was performed in Matlab 2017b (Mathworks). For the classification of clones, whole well mosaic images were segmented and classified with a custom algorithm: Briefly, the red channel was low pass filtered with a gaussian kernel of size 21 and standard deviation 7, and thresholded (RedPositiveMask = RedLP > 125). The green and red channels were low pass filtered with gaussian kernels of size 60 and standard deviation 20 and thresholded (GreenPositiveMask = GreenLP > 200, BluePositiveMask = BlueLP > 200). After segmentation, wells containing a clone with an area bigger than 20000 pixels were classified as Blue, Red, RedGreen, or Negative according to the area-proportion (AP) per channel. Area-proportion is defined as the count of pixels in a channel-specific mask divided by the count of pixels in the clonal region of interest, which is defined by the boolean OR operation between GreenPositiveMask and RedPositiveMask. Briefly, the class Red is defined as APred > 0.9 and APgreen < 0.1. The class Green is defined as APred < 0.1 and APgreen > 0.9, the class RedGreen is defined as APred > 0.1 and APgreen > 0.1, and all remaining clones are classified Negative.

The figure created for the graphical abstract was generated using images from Servier Medical Art Commons Attribution 3.0 Unported License. (http://smart.servier.com). Servier Medical Art by Servier is licensed under a Creative Commons Attribution 3.0 Unported License.

3. Results

3.1. Single-cell sorting of gene-edited iPSC clones

The pluripotent iPSC line A30P-4, (unique identifier HIHDNDi001-B) derived from the p.A30P *SNCA* patient underwent biallelic transfection containing the wild-type homology sequence attached with two different fluorescent reporters, a red fluorescent protein (dTOMATO) and a green fluorescent protein (EGFP). Biallelic editing using two fluorescent constructs provides the certainty of generating the homozygous wild-type isogenic cell line. Using FACS, double-discrimination for dead cells, cell clumps and doublets were used with strict gating (Figure 2A-C). The negative control used was the iPSC line without transfection (Figure 2D), a representative plot of the sorted $dTOMATO^{+}/EGFP^{+}$ cells is shown in Figure 2E. Outside of the homology arm there is a blue fluorescent protein (tagBFP) that allows for the identification of random integration events, which was removed by prior cell sorting (Supplementary Figure 1). The box and arrow in Figure 2E represents the gating used to sort the single-cells into the 96-well plate. Once the cells became visible, they were screened using the OPERA HCS with Matlab (version 2017b, Mathworks). The in-house developed image analysis algorithm (Supplementary Method 1) automates the segmentation of the cellular structure across three fluorescent channels and shows a merged image, with the clones numbered (Figure 2F). The use of HCS technology allows the operator to discriminate based on the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

presence of a single fluorescent construct only, a BFP^+ cell, iPS morphology and cell doublets from improper FACS gating.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

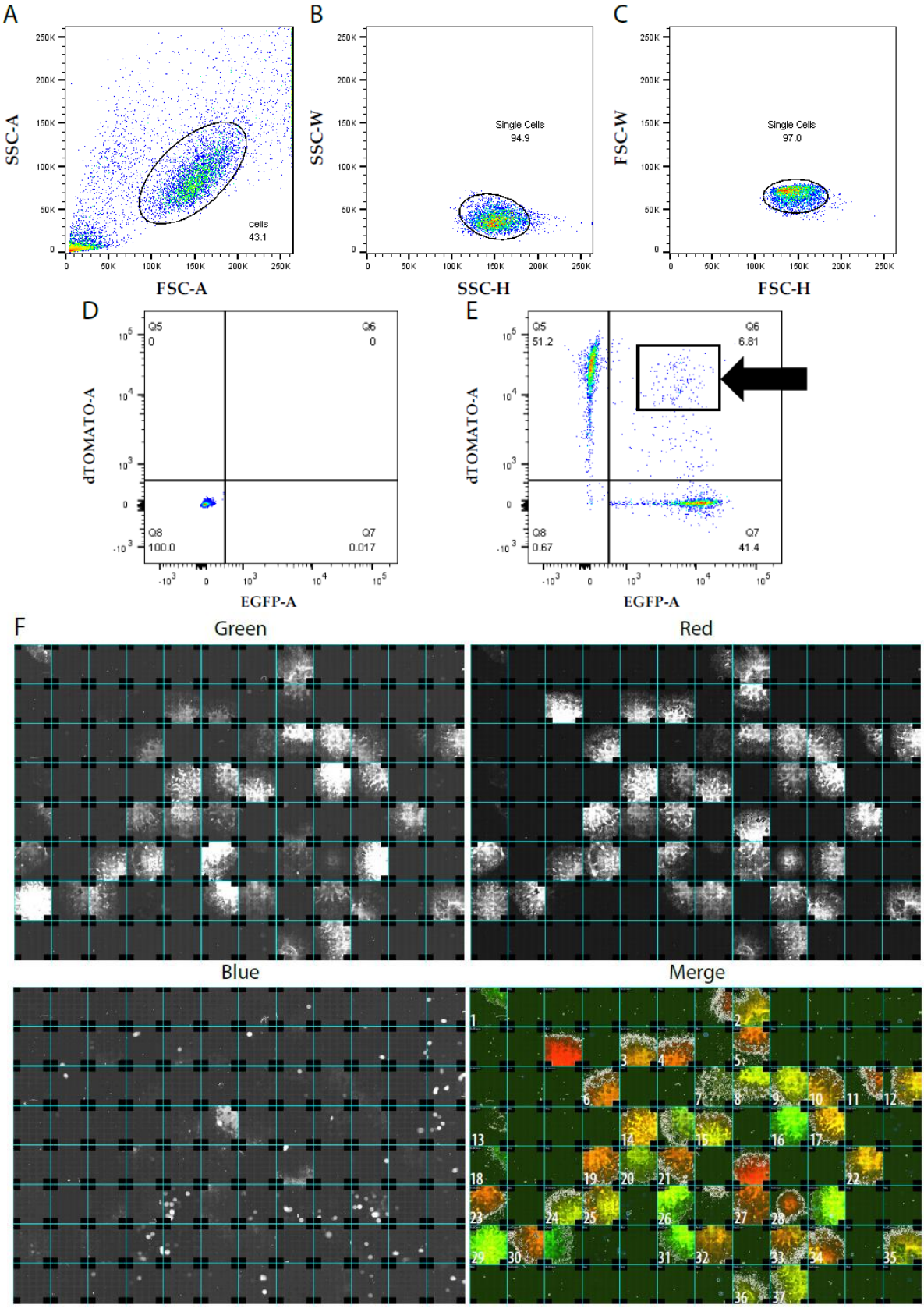

Figure 2: Sorting of single-cell isogenic iPSCs. **(A)** Selection of a live cell population and **(B-C)** doublet-discrimination **(D)** Generation of negative sorting gates using untransfected iPSCs. **(E)** Single-cell FACS sorting of dTOMATO^{+}/EGFP^{+} cells with restrictive gating (black arrow) into a 96 well plate. **(F)** High content screening of single-cell sorted plate showing the green, red, blue and merged channels.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

3.2. Restriction-digest mediated selection of single-cell clones

Following the positive selection and expansion of the monoclonal iPS cell lines using FACS and HCS, PCR amplification and restriction digestion was next set of selection criteria used to select the single-cell iPS cell lines (Figure 3). Of the 37 different monoclonal iPSC cell lines amplified (Figure 2F), 34 of those cell lines were successfully passaged and cryopreserved. Specific primers were used to amplify Exon 2 of the *SNCA* gene leading to a 314 bp product, a further 3 clones (Cl. 6, Cl. 14 and Cl.17) failed in the initial PCR amplification (Figure 3A), with Cl. 14 having introduced an insertion of the sequence (Supplementary Figure 2). The c.88G>C p.A30P *SNCA* heterozygous mutation leads to the generation of an *Mva1* restriction site. Conversely, if the heterozygous mutation has been repaired by HDR and gene-corrected, there should be neither the *Mva1* restriction site nor the digested product. Of the clones that are putatively isogenic, there are three distinct PCR products (Figure 3B). These are shown in Figure 3B as either a double-band (*), a lower band of approximately 100bp, or an undigested PCR product (#). Using Sanger sequencing (Figure 3C) we confirmed that the clones which had the double-band retained the heterozygous c.88G>C mutation and were not correctly edited. The iPS single-cell clones with the undigested PCR product at 314bp showed the successful generation of gene-corrected *SNCA* cell lines.

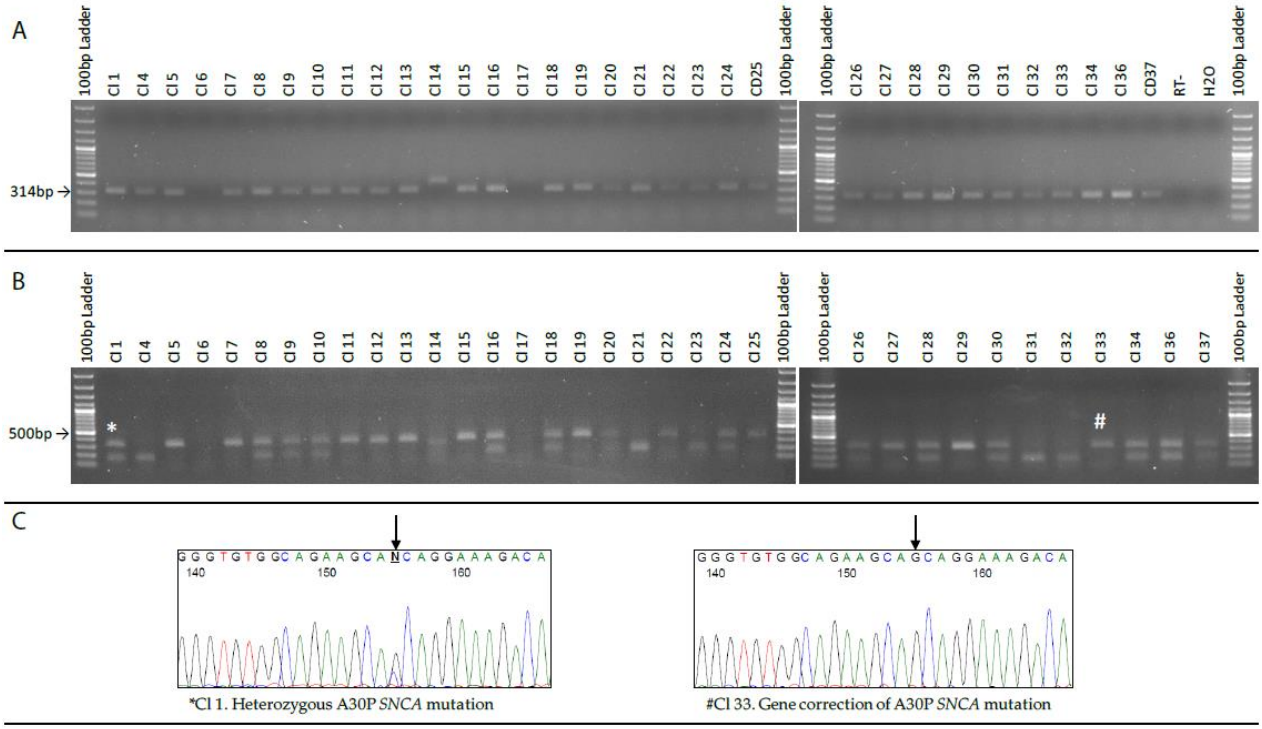

Figure 3: Restriction-digest mediated clonal selection. **(A)** PCR amplification of single-cell clones. **(B)** *Mva1* restriction digest post PCR amplification. *Double-band, #Unedited restriction digestion. **(C)** Sanger sequencing of amplified sequence. Black arrow signifies location of the c.88G>C *SNCA* mutation.

A list detailing the CRISPR-Cas9 mediated monoclonal isogenic patient-derived iPSC lines generated in this study, including cell-line validation criteria is shown in Table 1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

Table 1: List of CRISPR-Cas9 mediated single-cell isogenic patient-derived iPSC clones generated in this study with cell line validation criteria

Expanded single cell clones[1]	PCR Test	*Mva1* test digest	Sanger sequencing	Construct excision and genotyping
A30P-4 edited clone 1	314 bp	Digested: Double-band	Not done	
A30P-4 edited clone 4	314 bp	Digested: Lower bp product	Not done	
A30P-4 edited clone 5	314 bp	Undigested: 314 bp product	Isogenic: Gene-corrected p.A30P mutation	Chr 4: Deletion
A30P-4 edited clone 6	No product	No product	Not done	
A30P-4 edited clone 7	314 bp	Undigested: 314 bp product	Not done	
A30P-4 edited clone 8	314 bp	Digested: Double-band	Not done	
A30P-4 edited clone 9	314 bp	Digested: Double-band	Not done	
A30P-4 edited clone 10	314 bp	Digested: Double-band	Not done	
A30P-4 edited clone 11	314 bp	Undigested: 314 bp product	Not done	
A30P-4 edited clone 12	314 bp	Undigested: 314 bp product	Not done	
A30P-4 edited clone 13	314 bp	Undigested: 314 bp product	Isogenic: Gene-corrected p.A30P mutation	Normal genotype
A30P-4 edited clone 14	416 bp	Digested	102 bp insertion error	
A30P-4 edited clone 15	314 bp	Undigested: 314 bp product	Not done	
A30P-4 edited clone 16	314 bp	Digested: Double-band	Not done	
A30P-4 edited clone 17	No product	No product	102 bp insertion error	
A30P-4 edited clone 18	314 bp	Digested: Double-band	Not done	
A30P-4 edited clone 19	314 bp	Undigested: 314 bp product	Isogenic: Gene-corrected p.A30P mutation	
A30P-4 edited clone 20	314 bp	Digested: Double-band	Not edited: Heterozygous p.A30P *SNCA* mutation	
A30P-4 edited clone 21	314 bp	Digested: Lower bp product	Not done	
A30P-4 edited clone 22	314 bp	Undigested: 314 bp product	Not done	
A30P-4 edited clone 23	314 bp	Digested: Lower bp product	Not done	
A30P-4 edited clone 24	314 bp	Digested: Double-band	Not done	
A30P-4 edited clone 25	314 bp	Undigested: 314 bp product	Not done	
A30P-4 edited clone 26	314 bp	Digested: Double-band	Not done	
A30P-4 edited clone 27	314 bp	Undigested: 314 bp product	Not done	
A30P-4 edited clone 28	314 bp	Digested: Double-band	Not done	
A30P-4 edited clone 29	314 bp	Undigested: 314 bp product	Isogenic: Gene-corrected p.A30P mutation	
A30P-4 edited clone 30	314 bp	Digested: Double-band	Not done	

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

A30P-4 edited clone 31	314 bp	Digested: Lower bp product	Not done	
A30P-4 edited clone 32	314 bp	Digested: Lower bp product	Not done	
A30P-4 edited clone 33	314 bp	Undigested: 314 bp product	Isogenic: Gene-corrected p.A30P mutation	Normal genotype
A30P-4 edited clone 34	314 bp	Digested: Double-band	Not done	
A30P-4 edited clone 36	314 bp	Digested: Double-band	Not done	
A30P-4 edited clone 37	314 bp	Digested: Double-band	Not done	

[1] A30P-4 edited clones 2, 3 and 35 did not survive iPS passaging

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

3.3. Characterisation of single cell clones

Three cell lines of the gene-corrected isogenic cell lines (Cl. 5, Cl. 13, Cl. 33) were selected at random and transfected with an mRNA encoding the excision-only variant of the piggyBac transposase to remove the fluorescent constructs [16,26]. The subsequent dTOMATO^{-}/EGFP^{-} cell population were purified by three cell sorting steps (Supplementary Figure 3). The gene-corrected footprint-free isogenic cell lines were then karyotyped as part of the validation procedure. The gene-corrected cell lines Cl. 13 and Cl. 33 were karyotypically normal, had a stable genotype and passed this validation step, Cl. 5 had a deletion in Chromosome 4 and did not pass the validation step. All molecular karyotypes using single nucleotide polymorphism (SNP) analysis are shown in the supplementary materials (Supplementary figures 4-6).

The *de-novo* monoclonal gene-edited iPS cell lines were then characterised and validated by their pluripotency ability. The iPS colonies displayed a stem cell morphology and had positive marker staining for: OCT4, SOX2, NANOG and TRA-1-60 at the protein level (Figure 4A). The iPS cell lines also retained the *in-vitro* ability to directly differentiate into the three embryonic germ layers defined by specific marker expression specific to that lineage: Ectoderm (OTX2), Endoderm (SOX17) and Mesoderm (BRACHYURY) (Figure 4B).

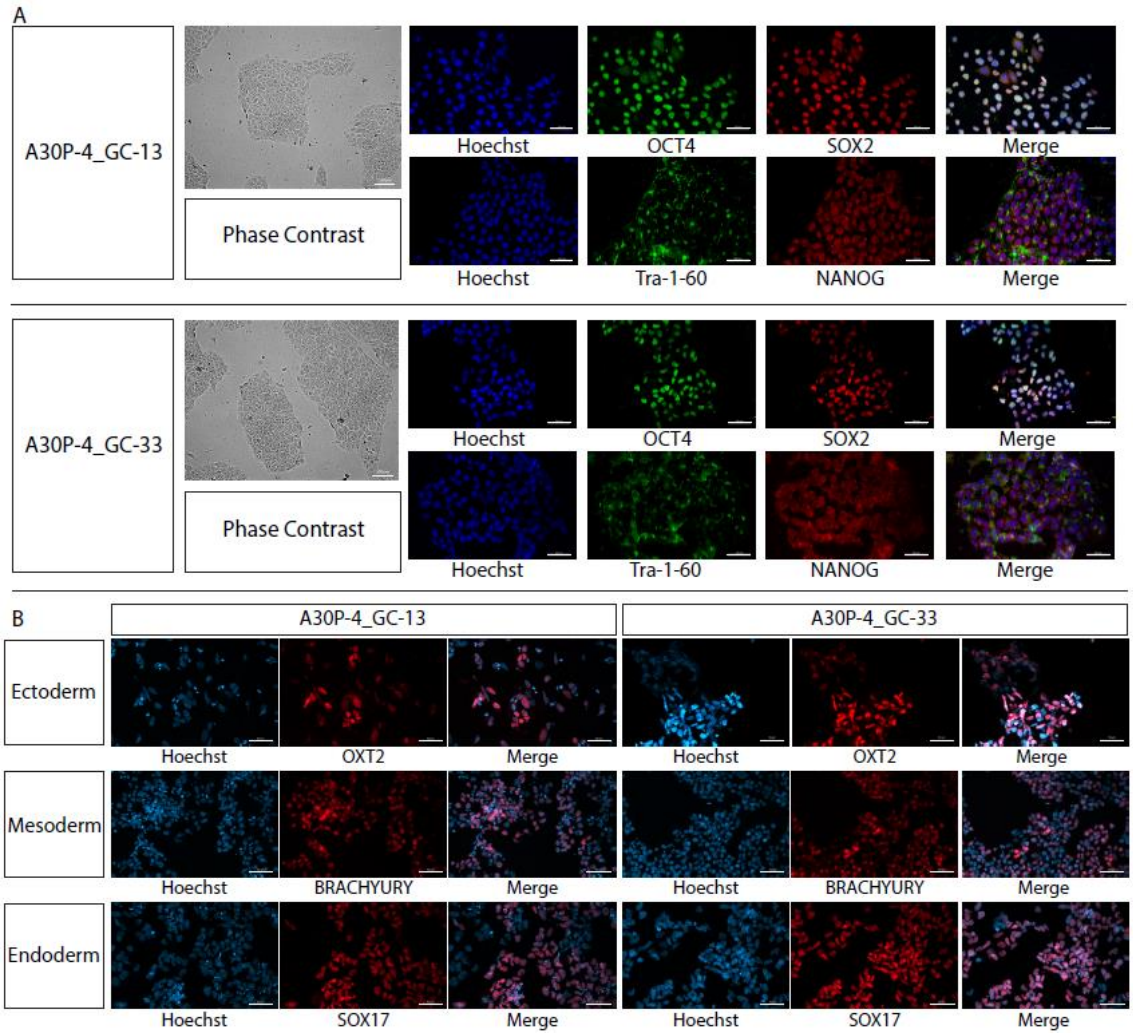

Figure 4: iPS cell line characterisation. **(A)** Pluripotency characterisation of the three single-cell isogenic lines generated in this project. Phase contrast images of pluripotent colonies, image taken using a 5x objective. Scale bar represents 100μM. Antibody specific pluripotency marker expression of OCT4/SOX2 and NANOG/TRA-1-60. Images taken using a 25x objective, scale bar is 50μM. **(B)** Directed differentiation to the three embryonic germ layers using antibodies specific to Ectoderm (OTX2), Endoderm (SOX17) and Mesoderm (BRACHYURY). Images taken using a 25x objective, scale bar is 50μM.

4. Discussion

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

This research article is the first to show gene-corrected isogenic cell lines from a PD patient carrying the mutation encoding p.A30P in the *SNCA* gene. Gene-correction of the patient iPSCs carrying the pathogenic p.A53T *SNCA* mutation and triplication of the *SNCA* gene locus have been previously published [27–29].

This study is the first to generate patient-derived gene-corrected single-cell isogenic iPS cell lines and is a novel approach within the literature. Furthermore, this study uses a quadruple selection strategy (QSS) to select, screen and validate these single-cell clones using genotyping and pluripotency characterisation. The use of the QSS: Antibiotic resistance; single-cell sorting and HCS; PCR amplification and restriction digest-mediated selection; and Sanger sequencing, to select clones is additionally a novel approach that has scope to be scaled up for high-throughput mediated generation of isogenic cell lines iPSCs. The use of the QSS will substantially reduce the time required to generate an isogenic cell line, in part by eliminating every other non-edited or incorrectly edited cell from the edited population. Sorting and screening multiple 96-well plates using HCS further reduces the time that the operator needs to spend manually checking and validating each colony/well through the microscope. PCR-mediated restriction digestion has long been used successfully in gene editing to validate generated patient-derived isogenic iPS clones [28] and introducing silent mutations or utilising the existing mutation within the sequence of the pre-edited cell line is one of the most effective strategies in validating the generated patient-derived isogenic iPS clone.

Using the QSS strategy to generate the single-cell isogenic clones gives an efficiency of 35.30% - 12 out of the 34 *Mva1* undigested clones were correctly edited (Table 2), which compares favourably to the standard isogenic generation rate of 1-5% [14,21]. However, it must be noted that without the sorting of the edited cells into single-cell clones, this efficiency would be markedly lower. The remaining 64.70% of sorted cells that had expressed both wild-type sequences shown by the presence of the respective dTOMATO and EGFP constructs were either incorrectly edited or unedited in the sequenced region. Without the single-cell isogenic method of selection, a heterozygous polyclonal mix of correctly and differentially edited cells would have been generated and not a bona-fide gene-corrected patient-derived isogenic iPS cell line. Moreover, the polyclonal cell line would not have passed the *Mva1* undigested restriction-digest validation step and the line would have not been generated, underlying the stringency of the validation.

Table 2: Summary of each stage of the gene-corrected p.A30P *SNCA* single-cell isogenic cell lines generated

	Sorted single-cell clones[1]	Expanded single-cell clones	PCR Amplification	*Mva1* undigested	Sequenced[2]	Normal Genotype[3]
Absolute numbers	37/192	34/37	31/34	12/34	5/5	2/3
Efficiency	19.27%	91.89%	91.18%	35.30%	100%	66.67%

[1] Only the yellow clones expressing both the dTOMATO$^+$/EGFP$^+$ constructs were amplified.

[2] Of the cell lines that passed the *Mva1* restriction digest criteria, five were chosen at random for Sanger sequencing.

[3] Of the five cell lines that were successfully sequenced, three were chosen at random for genotyping.

The plated recovery rate of the single-cell iPS clones is 19.27%, this already contains the correct fluorescent combination of dTOMATO$^+$/EGFP$^+$/tagBFP$^-$ (Table 2). A concern in the literature regarding the dissociation to single-cells and single-cell iPS culture is that there is a strong selective pressure for iPSCs to adapt that could lead to potential genomic abnormalities [30,31]. Changes in copy number variation (CNV) was used to sequence the genome and validate the isogenic cell lines after transposase-mediated excision of the construct. Two of the three single-cell isogenic cell lines selected at random had no genotypic abnormalities although the third did. However as this was due to a deletion on chromosome 4 (Table 1) it is likely this was mediated by gene editing as opposed to a cell culture acquired abnormality [32].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

5. Conclusion

The generation of gene corrected isogenic cell lines is critically important in understanding the mechanism how a pathogenic disease variant leads to the disease. Establishing a single-cell gene corrected isogenic as opposed to a heterozygous polyclonal isogenic is the next step in this research landscape. What we have shown is that the use of the QSS that includes the single-cell sorting and HCS easily allows the researcher to generate, validate and isolate the gene-corrected single-cell isogenic clones from the CRISPR-Cas9 correctly edited, incorrectly edited and non-edited heterozygous cellular population, improving on a technically challenging, variably efficient and time-consuming process.

Supplementary Materials: The following are available online. Figure S1: Removal of tagBFP^{+} cells using FACS prior to sorting the dTOMATO^{+}/EGFP^{+} transfected cells. Figure S2: 102 bp sequence insertion in Cl. 14. Figure S3: Transposase removal of biallelic constructs using FACS and validation of footprint-free isogenic cell lines. Figure S4: Molecular karyotype of the single-cell gene-corrected cell line clone 5. Figure S5: Molecular karyotype of the single-cell gene-corrected cell line clone 13. Figure S6: Molecular karyotype of the single-cell gene-corrected cell line clone 33. A frozen page containing the source code required for the implementation of the cloneclassifer script with the output shown in Figure 2F is available: https://doi.org/10.17881/lcsb.kcqg-tr55

Author Contributions: Conceptualization, P.B. (Peter Barbuti) and R.K. (Rejko Krüger); methodology, P.B. and G.C. (Gérald Cruciani); software, P.A. (Paul Antony); validation, P.B., B.S. (Bruno Santos), F.M. (Francois Massart) and C.D. (Claire Dording); formal analysis, P.B.; investigation, P.B.; resources, J.A. (Jonathan Arias) and J.S. (Jens Schwamborn); data curation, P.B. and P.A.; writing—original draft preparation P.B..; writing—review and editing, P.B., R.K., J.A., and P.A., visualization, P.B.; supervision, R.K.; project administration, P.B. and R.K.; funding acquisition, R.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by grants from the Fond National de Recherche within the PEARL programme (FNR/P13/6682797), the INTER programme (INTER/LEIR/18/12719318) and the National Centre for Excellence in Research on Parkinson's disease (NCER-PD) programme and by the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 692320 (WIDESPREAD; CENTRE-PD). The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Acknowledgments: The authors would like to acknowledge Laurent Heirendt and Yohan Jarosz of the Responsible and Reproducible Research (R3) initiative at the Bioinformatics core at the LCSB for uploading and verifying the software code.

Conflicts of Interest: J.A. and J.S. are inventors in a patent PCT/EP2017/051889. The remaining authors declare that the research was conducted in the absence of any financial or commercial relationships that could be construed as a potential conflict of interest.

References

1. Dorsey ER, Elbaz A, Nichols E, et al. Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol. 2018; 17:939–953

2. Spillantini MG, Crowther RA, Jakes R, et al. α-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with Lewy bodies. Proc. Natl. Acad. Sci. 1998; 95:

3. Polymeropoulos MH, Lavedan C, Leroy E, et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. Science 1997; 276:2045–7

4. Krüger R, Kuhn W, Müller T, et al. Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease. Nat. Genet. 1998; 18:106–108

5. Zarranz JJ, Alegre J, Gómez-Esteban JC, et al. The new mutation, E46K, of α-synuclein causes parkinson and Lewy body dementia. Ann. Neurol. 2004; 55:164–173

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

6. Lesage S, Anheim M, Letournel F, et al. G51D α-synuclein mutation causes a novel Parkinsonian-pyramidal syndrome. Ann. Neurol. 2013; 73:459–471

7. Pasanen P, Myllykangas L, Siitonen M, et al. Novel α-synuclein mutation A53E associated with atypical multiple system atrophy and Parkinson's disease-type pathology. Neurobiol. Aging 2014; 35:2180.e1–5

8. Chartier-Harlin M-C, Kachergus J, Roumier C, et al. Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. Lancet (London, England) 2004; 364:1167–9

9. Singleton AB, Farrer M, Johnson J, et al. alpha-Synuclein locus triplication causes Parkinson's disease. Science 2003; 302:841

10. Campêlo CL das C, Silva RH. Genetic Variants in SNCA and the Risk of Sporadic Parkinson's Disease and Clinical Outcomes: A Review. Parkinsons. Dis. 2017; 2017:1–11

11. Simón-Sánchez J, Schulte C, Bras JM, et al. Genome-wide association study reveals genetic risk underlying Parkinson's disease. Nat. Genet. 2009; 41:1308–1312

12. Mali P, Yang L, Esvelt KM, et al. RNA-Guided Human Genome Engineering via Cas9. Science (80-.). 2013; 339:823 LP – 826

13. Ran FA, Hsu PD, Wright J, et al. Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. 2013; 8:2281–2308

14. Grobarczyk B, Franco B, Hanon K, et al. Generation of Isogenic Human iPS Cell Line Precisely Corrected by Genome Editing Using the CRISPR/Cas9 System. Stem Cell Rev. Reports 2015; 11:774–787

15. Flynn R, Grundmann A, Renz P, et al. CRISPR-mediated genotypic and phenotypic correction of a chronic granulomatous disease mutation in human iPS cells. Exp. Hematol. 2015; 43:838-848.e3

16. Arias-Fuenzalida J, Jarazo J, Qing X, et al. FACS-Assisted CRISPR-Cas9 Genome Editing Facilitates Parkinson's Disease Modeling. Stem Cell Reports 2017; 9:1423–1431

17. Gomez-Giro G, Arias-Fuenzalida J, Jarazo J, et al. Synapse alterations precede neuronal damage and storage pathology in a human cerebral organoid model of CLN3-juvenile neuronal ceroid lipofuscinosis. Acta Neuropathol. Commun. 2019; 7:222

18. Hanss Z, Boussaad I, Jarazo J, et al. Quality Control Strategy for CRISPR-Cas9-Based Gene Editing Complicated by a Pseudogene . Front. Genet. 2020; 10:1297

19. Yang R, Lemaître V, Huang C, et al. Monoclonal Cell Line Generation and CRISPR/Cas9 Manipulation via Single-Cell Electroporation. Small 2018; 14:1702495

20. Xu X, Tay Y, Sim B, et al. Reversal of Phenotypic Abnormalities by CRISPR/Cas9-Mediated Gene Correction in Huntington Disease Patient-Derived Induced Pluripotent Stem Cells. Stem Cell Reports 2017; 8:619–633

21. Wang G, Yang L, Grishin D, et al. Efficient, footprint-free human iPSC genome editing by consolidation of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2020 doi:10.20944/preprints202005.0001.v1

Cas9/CRISPR and piggyBac technologies. Nat. Protoc. 2017; 12:88–103

22. Krüger R, Kuhn W, Leenders KL, et al. Familial parkinsonism with synuclein pathology: clinical and PET studies of A30P mutation carriers. Neurology 2001; 56:1355–62

23. Barbuti PA, Santos BFR, Dording CM, et al. Generation of two iPS cell lines (HIHDNDi001-A and HIHDNDi001-B) from a Parkinson's disease patient carrying the heterozygous p.A30P mutation in SNCA. Manuscr. Submitt. Publ. 2020;

24. Larsen SB, Hanss Z, Cruciani G, et al. Induced pluripotent stem cell line (LCSBi001-A) derived from a patient with Parkinson's disease carrying the p.D620N mutation in VPS35. Stem Cell Res. 2020; 101776

25. Peiffer DA, Le JM, Steemers FJ, et al. High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. Genome Res. 2006; 16:1136–1148

26. Li X, Burnight ER, Cooney AL, et al. piggyBac transposase tools for genome engineering. Proc. Natl. Acad. Sci. 2013; 110:E2279–E2287

27. Ryan SD, Dolatabadi N, Chan SF, et al. Isogenic Human iPSC Parkinson's Model Shows Nitrosative Stress-Induced Dysfunction in MEF2-PGC1α Transcription. Cell 2013; 155:1351–1364

28. Soldner F, Laganière J, Cheng AW, et al. Generation of Isogenic Pluripotent Stem Cells Differing Exclusively at Two Early Onset Parkinson Point Mutations. Cell 2011; 146:318–331

29. Heman-Ackah SM, Manzano R, Hoozemans JJM, et al. Alpha-synuclein induces the unfolded protein response in Parkinson's disease SNCA triplication iPSC-derived neurons. Hum. Mol. Genet. 2017; 26:4441–4450

30. Peterson SE, Loring JF. Genomic Instability in Pluripotent Stem Cells: Implications for Clinical Applications. J. Biol. Chem. 2014; 289:4578–4584

31. Bai Q, Ramirez J-M, Becker F, et al. Temporal Analysis of Genome Alterations Induced by Single-Cell Passaging in Human Embryonic Stem Cells. Stem Cells Dev. 2014; 24:653–662

32. Mayshar Y, Ben-David U, Lavon N, et al. Identification and Classification of Chromosomal Aberrations in Human Induced Pluripotent Stem Cells. Cell Stem Cell 2010; 7:521–531