bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Organoid Easytag: an efficient workflow for gene targeting in human organoids

Dawei Sun[1,2], Lewis D. Evans[3] and Emma L. Rawlins[1,4]

[1] Wellcome Trust/CRUK Gurdon Institute, University of Cambridge, Cambridge, CB2 1QN, UK.
[2] Department of Pathology, University of Cambridge, Cambridge, CB2 1QP, UK.
[3] Developmental Biology and Cancer Department, University College London, London, WC1N 1DZ, UK.
[4] Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3EL, UK.

Contact: e.rawlins@gurdon.cam.ac.uk

Key words: human, organoids, CRISPR-Cas9, homologous gene targeting.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Human organoid systems recapitulate key features of organs offering platforms for modelling human developmental biology and disease. However, tissue-derived organoids suffer from low efficiency of genetic manipulations, especially CRISPR/Cas9-mediated knock-ins. We have systematically optimised and developed an "Organoid Easytag" pipeline for efficient (40-65%) and accurate gene targeting to facilitate generation of reporter lines and gene knock-outs in organoid based research.

Main

Precise targeted gene integration has been established in organoids, but is not efficient. A recent application of non-homologous end joining (NHEJ) to improve gene targeting in organoids has been successful[1]. We present a complementary approach to achieve efficient gene targeting using a homology directed repair (HDR) strategy: Organoid Easytag. Organoid Easytag uses Cas9 protein and synthetic gRNA ribonucleoprotein (RNP), a circular repair template and fluorescence-activated cell sorting to enrich targeted cells. The recombination-based approach means that our strategy has the power to deliver precise genetic manipulations with minimal additional changes. We developed the Organoid Easytag workflow (Fig. 1a) in human foetal lung organoids[2]. The pipeline can be easily adapted for other organoid systems.

Gene targeting methods were largely developed for pluripotent stem cells and typically use a drug selection cassette and plasmid-encoded-Cas9 which, in organoids, results in chimeric colonies and low editing efficiency, respectively[3,4]. We chose fluorescence as a selection marker, allowing targeted cells to be easily isolated using flow cytometry and any chimeric colonies to be identified and removed using a fluorescent microscope.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In order to achieve efficient gene targeting, we first sought to maximise: (1) the efficiency of DNA delivery into organoid cells, which impacts repair template presentation; (2) the efficiency of site-specific DNA cleavage by the Cas9-gRNA complex, which influences the likelihood of homology directed repair (HDR) being triggered. Commonly used methods for mammalian cell DNA delivery, including lipofectamine, nucleofection and lentivirus, were tested. Nucleofection consistently achieved up to 70% transfection efficiency across different organoid lines (Supplementary Fig. 1a, 1b). To optimize site-specific DNA cleavage, we used nucleofection to introduce the Cas9-gRNA complex into cells in different forms (Supplementary Fig. 1c): (A) RNP complex formed of purified Cas9 protein and synthetic crispr/tracer gRNA heterodimer (cr/tr RNP); (B) RNP of Cas9 protein and synthetic single strand gRNA (ssRNP); (C) plasmids encoding Cas9 protein and gRNA. Consistent with previous reports, Cas9 RNPs out-performed plasmid based Cas9 approaches[5,6] both in the T7 endonuclease assay and an online CRISPR editing analysis tool (Supplementary Fig. 1c, 1d). Remarkably, the ssRNP generated almost twice the amount of indels compared to the cr/tr RNP. Thus, we adopted nucleofection and ssRNP for downstream experiments. This strategy has the advantage that the RNP is rapidly degraded and should produce minimal off-target effects.

To optimize our workflow (Fig. 1a), we first focused on generating an *ACTB*-fusion protein, taking advantage of the abundance of ACTB protein in human foetal lung organoids and a previously-published *ACTB* targeting strategy[7]. Having an efficient gRNA for the *ACTB* locus (Supplementary Fig. 1c), we designed a repair template to generate an N terminal monomeric (m)*EGFP-ACTB* fusion (Fig. 1b). We set the following rules for repair template design to facilitate efficient and consistent gene targeting: (1) protospacer adjacent motif (PAM) sequence mutated to prevent editing by ssRNP[8]; (2) 700 to 1000

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

nucleotide length of each homologous arm[9]; (3) minimal plasmid size to maximise delivery into organoid cells; (4) monomeric forms of fluorescent protein to avoid undesirable fusion protein aggregates. As expected, 72 hours after nucleofection of the ssRNP and repair template, mEGFP$^+$ organoid cells could be enriched by flow cytometry (Fig. 1c). These cells were collected and pooled together, but seeded sparsely, and were successfully expanded into organoid colonies (Fig. 1d). The mEGFP-ACTB fusion protein localized to cell–cell junctions, consistent with previous reports[7]. These small colonies could be further expanded into new organoid lines and 59% of lines (n = 17/29 lines, from N = 2 parental organoid lines) were correctly targeted. Targeted organoids continued to express the multipotent lung progenitor marker, SOX9 (Fig. 1e).

We sought to further increase targeting efficiency (Supplementary Fig. 2a). Previous research reported that various drugs including a RAD51 agonist (RS-1), a β3-adrenergic receptor agonist (L755507) and a DNA ligase IV inhibitor (SCR-7) enhance gene targeting efficiency[10-13]. However, using flow cytometry as a simple assay, none of the drugs tested increased the rate of gene targeting in the organoids (Supplementary Fig. 2b).

Synthetic single-stranded donor oligonucleotide (ssODN) repair templates offer a ‘cloning free’ workflow for targeting short peptide tags into different loci[14-16], although they are limited by oligonucleotide length. We explored ssODN performance in our Easytag workflow using the split GFP system, by tagging the N terminal of *ACTB* with *GFP11* (Supplementary Fig. 3a). We provided *GFP11-ACTB* ssODN, or a positive-control plasmid, repair template together with a small transient GFP(1-10) expressing vector. Targeted cells would become GFP$^+$ in the presence of the GFP(1-10) vector. We obtained GFP$^+$ cells by flow cytometry in both the plasmid and ssODN repair template groups

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Supplementary Fig. 3b). However, when ssODN was used no organoid colonies formed after 10 days of culture, whereas the plasmid group started to show numerous colonies (Supplementary Fig. 3c). We reasoned that ssODN could be error prone and generate indels in the *ACTB* locus, which would be detrimental to cells. We tested this hypothesis by tagging SOX2, a protein that cells are less sensitive to, with a V5 tag using an ssODN repair template (Supplementary Fig. 3d). In this experiment, we obtained $V5^+$ colonies with low efficiency (1/14 colonies with V5 integration in *SOX2*) and the one tagged allele obtained had a random insertion near the gRNA cutting site (Supplementary Fig. 3e, 3e'). We conclude that synthetic ssODN templates are error prone and not optimal for the organoid Easytag workflow.

To expand our pipeline to target other loci, we targeted *SOX9*, a transcription factor, to represent genes expressed in a less abundant manner. SOX9 is a tip progenitor cell marker for developing lungs[2]. Thus, incorrect targeting may have a detrimental impact and *SOX9* reporters are useful for monitoring progenitor state in human foetal lung organoids. In order to overcome the low expression level of *SOX9*, we used a Histone H2B-EGFP fusion (H2B-EGFP hereafter) to concentrate the EGFP signal in nucleus (Fig. 2a). A *T2A* sequence, a self-cleavage peptide, was also inserted between *SOX9* and *H2B-EGFP*, to ensure that SOX9 protein was minimally influenced. This strategy allowed us to enrich correctly targeted cells. Colonies could be expanded and maintained normal SOX2, SOX9 and NKX2.1 expression (Fig. 2b, Supplementary Fig. 4a). Importantly, we noted that although we were only able to generate *SOX9* reporter lines as heterozygotes (Supplementary Fig 4b, 4c), the gRNA sites in the wildtype alleles were intact (6/6 lines tested, N=3 parental organoid lines) (Supplementary Fig. 4d). This offers the opportunity of retargeting the second allele if desired.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The *AAVS1* locus has been generally considered to be 'safe harbour locus' for expressing exogenous genes in a controllable manner in human cells without silencing[17]. We sought to target the *AAVS1* locus in organoids. As a proof of concept, we targeted a membrane tagged TagRFP-T (mTagRFP-T) to visualise cell shapes under the control of an EF1α promoter to the *AAVS1* locus (Fig. 2c). Using the Organoid Easytag workflow, we generated organoid colonies with mTagRFP-T labelled cell membranes (Fig. 2d). Overall, our gene targeting efficiency was 40-65% across different organoid lines (Fig. 2e).

Generation of straightforward gene knockouts using the CRISPR-Cas9 system can suffer from translation retention and exon skipping[18,19]. Moreover, in the absence of a strong, immediate phenotype the knockout cells cannot readily be identified. We sought to solve these problems by generating a gene knockout in a more controlled manner using the Organoid Easytag workflow. We focused on the *SOX2* gene as its function remains to be elucidated in human foetal lung progenitors. We swapped the *SOX2* exon with *T2A-H2B-EGFP* to generate *SOX2* knockout organoids. Using two gRNAs targeting the N and C terminal of the *SOX2* coding sequence respectively, we sequentially replaced both copies of the *SOX2* coding sequence (CDS) (Figure 2f, 2g; Supplementary Fig. 5). This again illustrated the power of re-targeting the second allele using the Organoid Easytag workflow. *SOX2* knockout colonies can proliferate and grow normally, suggesting that SOX2 is not crucial for human foetal lung tip progenitor cell self-renewal. Thus, we have generalised our Organoid Easytag pipeline to target various loci, including highly abundant genes, transcription factors, the human safe harbour locus, and to generate knockouts.

The Organoid Easytag workflow provides a robust and versatile pipeline to perform gene targeting in the human foetal lung organoid system. Heterozygous knock-ins can be efficiently generated in organoids to produce reporter lines.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Moreover, the WT allele typically remains intact, providing the option of targeting the second allele which enables generation of knockouts. We envision that the Organoid Easytag workflow can be easily adapted for gene targeting in other systems, including organoids from other tissues, human pluripotent stem cells, cancer cell lines and cancer organoids.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Methods

Derivation and maintenance of human foetal lung organoid culture

Human foetal lung organoids were derived and maintained as previously reported (Nikolic et al., 2017). Briefly, human foetal lung tissues were dissociated using Dispase (8 U/ml Thermo Fisher Scientific, 17105041) at room temperature (RT) for 2 min. Mesenchyme was dissected away using needles. Tips of the branching epithelium were micro-dissected, transferred into 50 µl of Matrigel (356231, Corning) and seeded in one well of a 24 well low-attachment plate (M9312-100EA, Greiner). The plate was incubated at 37°C for 5 min to solidify the Matrigel. 600 µl of self-renewing medium containing: N2 (1: 100), B27 (1: 50), N-acetylcysteine (1.25 mM), EGF (50 ng/ml, PeproTech, AF-100-15), FGF10 (100 ng/ml, PeproTech, 100-26), FGF7 (100 ng/ml, PeproTech, 100-19), Noggin (100 ng/ml, PeproTech, 120-10C), R-spondin-1 conditioned medium (5% v/v, Stem Cell Institute, University of Cambridge), CHIR99021 (3 µM, Stem Cell Institute, University of Cambridge) and SB-431542 (10 µM, bio-techne, 1614), was added. The plate was incubated under standard tissue culture conditions (37°C, 5% CO_2). Once formed, organoids were maintained in self-renewing medium and passaged by mechanically breaking using P200 pipettes every 10-14 days.

Whole mount immunostaining for human foetal lung organoids

Organoids were fixed with 4% paraformaldehyde (PFA) directly in the culture plates on ice for 30 min. After two PBS washes, 0.5% (w/v) Bovine Serum Albumin (BSA), 0.2% Triton-X in PBS (washing solution) was added and left on ice overnight to dissolve Matrigel. Organoids were then transferred into multiple CellCarrier-96 Ultra Microplates (PerkinElmer, 6055300) for staining.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Subsequently, blocking was performed in 5% donkey serum (Stratech, 017-000-121-JIR), 0.5% (w/v) Bovine Serum Albumin (BSA), 0.2% Triton-X in PBS (blocking solution) at 4°C overnight. For primary antibody staining, the following antibodies in blocking solution were used at 4°C overnight: SOX2 (1: 500, Bio-techne, AF2018), SOX9 (1: 500, Sigma, AB5535), E-cadherin (1: 1500, Thermo Fisher Scientific, 13-1900), NKX2-1 (1: 500, AbCam, ab76013), TagRFP (1: 1000, Evrogen, AB233), GFP (1: 500, AbCam, ab13970). After washing off the primaries, the following secondary antibodies in washing buffer were used at 4°C overnight: donkey anti-chick Alexa 488 (1: 2000, Jackson Immune, 703-545-155), donkey anti-rabbit Alexa 594 (1: 2000, Thermo Fisher Scientific, A-21207), donkey anti-goat Alexa 594 (1: 2000, Thermo Fisher Scientific, A-11058), donkey anti-goat Alexa 647 (1: 2000, Thermo Fisher Scientific, A-21447), donkey anti-rat Alexa 647 (1: 2000, Jackson Immune, 712-605-153). The following day, DAPI (Sigma, D9542) staining was performed in washing solution at 4°C for 30 min. After two washes with PBS, 97% (v/v) 2’−2’-thio-diethanol (TDE, Sigma, 166782) in PBS was used for mounting. Confocal z stacks were acquired using Leica SP8 at an optical resolution of 1024 × 1024 at 40x. Single z plane images are shown. Images were processed using ImageJ (version 2.0.0).

Plasmid nucleofection

For testing transduction efficiency, organoids were dissociated into single cells using pre-warmed TrypLE™ Express (12605028, Thermo Fisher Scientific) at 37°C for 10 min. The reaction was terminated by adding Advanced DMEM/F12 (12634028, Thermo Fisher Scientific) and cells passed through a 30-micron cell strainer. 2×10^5 organoid single cells were re-suspended with Lonza P3 nucleofection buffer and 1 μl of pmaxGFP (Lonza) and transferred to a 20 μl nucleofection cuvette (V4XP-3024, Lonza). Nucleofection was performed using

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Lonza 4D Nucleofector with X unit using program EA125. After nucleofection, self-renewing medium supplemented with 10 μM Y-27632 (ROCK inhibitor, ROCKi, 688000, Merck) was added to dilute the P3 buffer. Cell mixture was then seeded in Matrigel in 2 wells of a 24-well plate and cultured with self-renewing medium with ROCKi (10 μM) for 72 hrs before FACS analysis.

Lentiviral production

We grew HEK293T cells in 10-cm dishes to a confluence of 80% before transfecting the lentiviral vector (10 μg) with packaging vectors including pMD2.G (3 μg, Addgene plasmid # 12259), psPAX2 (6 μg, Addgene plasmid # 12260) and pAdVAntage (3 μg, E1711, Promega) using Lipofectamine 2000 Transfection Reagent (11668019, Thermo Fisher Scientific) according to manufacturer's protocol. After 16 hrs, medium was refreshed. Supernatant containing lentivirus was harvested at 24 hrs and 48 hrs and pooled together. Lentivirus was then concentrated using Lenti-X™ Concentrator (631232, Takara). Lentivirus pellets were dissolved in 200 μl PBS.

Lentivirus infection of organoids

Organoid single cell suspension was prepared as for nucleofection. 5 μl lentivirus (CMV-myrAKT-IRES-GFP) suspension was applied to 2×10^5 organoid single cells suspended in 500 μl self-renewing medium with 10 μM ROCKi (without Matrigel) in one well of 24-well plate and incubated at 37 °C overnight. The following day, cells were harvested and washed twice with PBS before pelleting and seeding, in Matrigel, in two wells of 24-well plate. Cells were grown in self-renewing medium with ROCKi (10 μM) for 72 hrs before flow cytometry. CMV-myrAKT-IRES-GFP was used for checking lentiviral transduction efficiency.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Lipofectamine transfection of organoids

Organoid single cells were prepared the same way as for nucleofection. For comparing transduction efficiency, 1 µg of pmaxGFP (Lonza) was mixed with 1 µl of Lipofectamine™ Stem Transfection Reagent (STEM00001, Thermo Fisher Scientific) according to manufacturer's protocol. 50 µl reaction mixture was applied to 2×10^5 organoid single cells suspended with 450 µl self-renewing medium with ROCKi (without Matrigel) in a single well of a 24-well plate. The plate was then centrifuged at 32°C at 600g for 1 hr, followed by incubation at 37°C for 2-4 hrs. Cells were then harvested, pelleted and seeded in Matrigel in two wells of a 24-well plate and grown in self-renewing medium supplemented with ROCKi (10 µM) for 72 hrs before FACS analysis.

Nucleofection for gene targeting

Cas9 protein was prepared and used as previously reported (Bruntraeger et al., 2019). If synthetic single strand gRNAs were used, 2 µl spCas9 (4 µg/µl) and 2.88 µl of ssRNA (100 µM, Synthego) were mixed and incubated at RT for a minimum of 10 min in order to form ssRNPs. If synthetic cr/tr RNA heterodimers were used, 200 pmol synthetic cr RNA (IDT) and 200 pmol synthetic tr RNA (IDT) were mixed with 2.5 µl Nuclease Free Duplex Buffer (11-01-03-01, IDT) and denatured at 95 °C for 2 min. 2 µl of cr/tr RNA heterodimer was cooled down to RT on the bench, mixed with 2 µl spCas9 (4 ug/ul) and incubated at RT for a minimum of 10 min to form cr/tr RNPs. At the same time organoids were dissociated into single cells, according to the protocol previously described for nucleofection. 4×10^5 cells were suspended using Lonza Nucleofection P3 buffer, mixed with 10 µg of appropriate plasmid repair template, or with 500 pmol ACTB-GFP11 ssODN (Ultramer DNA Oligos, IDT)

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

with 5 µg CMV-GFP(1-10) plasmid. The cell suspension was further mixed with pre-formed Cas9 RNPs and equally distributed into two 20 µl cuvettes (V4XP-3024, Lonza). Nucleofection was performed using program EA104. After nucleofection, self-renewing medium with ROCKi was added to dilute the P3 buffer. Cell mixture was then taken out and seeded in Matrigel in 4 wells of a 24-well plate and cultured with self-renewing medium with 10 µM ROCKi for 72 hrs before flow cytometry.

gRNA sequences used are as follows:
ACTB 5'-GCTATTCTCGCAGCTCACCA TGG,
SOX9 5'-CTTGAGGAGGCCTCCCACGA AGG,
AAVS1 5'-GTCCCCTCCACCCCACAGTG GGG,
SOX2 N terminal 5'-CGGGCCCGCAGCAAACTTCG GGG,
SOX2 C terminal 5'-CGGCCCTCACATGTGTGAGA GGG.
PAM sequences are underlined.

ACTB-GFP11 ssODN sequence is as follows:
5'*CGCCGGAAGTGGCCAGGGCGGGGGCGACCTCGGCTCACAGCGCGCC CGGCTATTCTCGCAGCTCACCATGCGTGACCACATGGTCCTTCATGAGTA TGTAAATGCTGCTGGGATTACAGCCGGCTCCGGTACCGATGATGATATCG CCGCGCTCGTCGTCGACAACGGCTCCGGCATGTGCAAGGCCGGCTTCG CGGGCG*-3'

SOX2-V5 ssODN sequence is as follows:
5'*TTCACATGTCCCAGCACTACCAGAGCGGCCCGGTGCCCGGCACGGCC ATTAACGGCACACTGCCGCTCTCACACATGGGTAAGCCTATCCCTAACCC TCTCCTCGGTCTCGATTCTACGTGAGGGCCGGACAGCGAACTGGAGGGG GGAGAAATTTTCAAAGAAAAACGAGGGAAATGGGAGGGGTGCAAAAGAG GAGA*-3'

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Small molecule influence on gene targeting efficiency

mEGFP-ACTB gene targeting was performed as previously described. After nucleofection, DMSO (0.6 ul, D2650, Sigma), RS-1 (10 uM, R9782, Sigma), L755507 (5 uM, SML1362, Sigma), or SCR-7 (100 uM, SML1546, Sigma) were added to self-renewing medium with ROCKi for 48 hrs. Organoid cells were analysed by flow cytometry 72 hrs after nucleofection.

T7 Endonuclease Assay

To test for site specific DNA cleavage using the T7 endonuclease assay, organoid cells were harvested 48 hrs after nucleofection of ssRNP, tr/cr RNP, plasmid encoding Cas9 and gRNA or WT control organoids. Genomic DNA was extracted using QIAamp Fast DNA Tissue Kit (51404, Qiagen). PCR was performed using PrimeSTAR® GXL DNA Polymerase (R050A, Takara) with 20 ng of genomic DNA as template according to manufacturer's protocol. Forward primer: 5'-TTGCCAATGGGGATCGCAG-3' and reverse primer: 5'-GCTCGATGGGGTACTTCAGG-3' were used for *ACTB* locus amplification. 10 µl of PCR product was then mixed with 1.5 µl 10X NEBuffer 2 (B7002S, NEB) and 1.5 µl of Nuclease-free water. The mixture was denatured at 95°C for 10 min, followed by ramp -2°C per second from 95°C to 85°C and ramp -0.3 °C per second from 85°C to 25°C. 2 µl T7 Endonuclease I (1 U/ul, M0302S, NEB) was added and incubated at 37°C for 1 hr. 2.5% Agarose gel was used for electrophoresis.

ICE analysis for Indel production

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Genomic DNA was extracted from organoid cells were harvested 48 hrs after nucleofection of ssRNP, tr/cr RNP, plasmid encoding Cas9 and gRNA or WT control organoids using QIAamp Fast DNA Tissue Kit (51404, Qiagen). PCR was performed using PrimeSTAR® GXL DNA Polymerase (R050A, Takara) with 20 ng of genomic DNA as template according to manufacturer's protocol. Forward primer: 5'-TTGCCAATGGGGATCGCAG-3' and reverse primer: 5'-GCTCGATGGGGTACTTCAGG-3' were used for *ACTB* locus amplification. PCR products were cleaned up using Macherey-Nagel™ NucleoSpin™ Gel and PCR Clean-up Kit (Macherey-Nagel, 740609.50) and sent for Sanger Sequencing (Department of Biochemistry, University of Cambridge) using reverse primer: 5'-GCTCGATGGGGTACTTCAGG-3'. Sanger sequencing results were compared using ICE online CRISPR editing analysis tool: https://www.synthego.com/products/bioinformatics/crispr-analysis. Statistical tests were two-tailed Student's t-tests.

Flow cytometry analysis

Organoid single cells were prepared 72 hrs after nucleofection, lentivirus transduction or Lipofectamine transfection. Cells were analysed using Sony SH800S Cell Sorter and Flowjo software (version 10.4).

Plasmid Construction

eSpCas9(1.1) was a gift from Feng Zhang (Addgene plasmid # 71814). *ACTB* gRNA sequence *5'-GCTATTCTCGCAGCTCACC-3'* was cloned into the vector using BbsI sites. *ACTB* repair template AICSDP-15: ACTB-mEGFP was a gift from The Allen Institute for Cell Science (Addgene plasmid # 87425). *SOX9* repair template was created by Infusion (638909, Takara) cloning to insert *SOX9* 5' and 3' homologous arms in EasyFusion T2A-H2B-GFP plasmid (a gift

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

from Janet Rossant, Addgene plasmid # 112851). *AAVS1* repair template was created by Infusion cloning to swap the CAG promoter and Puromycin resistance cassette in plasmid AICSDP-42: AAVS1-mTagRFPT-CAAX (a gift from The Allen Institute for Cell Science, Addgene plasmid # 107580). *SOX2* knockout repair template was created by Infusion cloning to insert *SOX2* 5' and 3' homologous arms in EasyFusion T2A-H2B-GFP (a gift from Janet Rossant, Addgene plasmid # 112851). The minimal CMV-GFP(1-10) plasmid was created by Infusion cloning of CMV-GFP(1-10) from pcDNA3.1-GFP(1-10) (a gift from Bo Huang (Addgene plasmid # 70219) into a pUC57 backbone. For testing lentiviral transduction efficiency, CMV-myrAKT-IRES-GFP vector was created by Infusion cloning to insert myrAKT from pCCL-Akt1 (a gift from Bi-Sen Ding, Icahn School of Medicine, Mount Sinai) and IRES sequence into pFP945 (a gift from Frederick Livesey, University College London). All plasmids will be deposited in Addgene.

Organoid Genotyping

Organoids from a single 48-well plate well were used for genomic DNA extraction with QIAamp Fast DNA Tissue Kit (51404, Qiagen) according to manufacturer's protocol. PCR was performed using PrimeSTAR® GXL DNA Polymerase (R050A, Takara) with 20 ng of genomic DNA as template according to manufacturer's protocol. Primers used are listed in genotyping primers file. For each gene targeting, 3 randomly picked lines were chosen for further Sanger Sequencing. 5' and 3' homologous arms of the gene targeting product were amplified using PrimeSTAR® GXL DNA Polymerase with aforementioned primers. PCR products were cleaned up using Macherey-Nagel™ NucleoSpin™ Gel and PCR Clean-up Kit (Macherey-Nagel, 740609.50) and sequenced using Sanger Sequencing (Department of Biochemistry, University of Cambridge).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. Artegiani, B. *et al.* Fast and efficient generation of knock-in human organoids using homology-independent CRISPR-Cas9 precision genome editing. *Nat. Cell Biol.* **22,** 321–331 (2020).
2. Nikolić, M. Z. *et al.* Human embryonic lung epithelial tips are multipotent progenitors that can be expanded in vitro as long-term self-renewing organoids. *Elife* **6,** e26575 (2017).
3. Fujii, M. *et al.* Human Intestinal Organoids Maintain Self-Renewal Capacity and Cellular Diversity in Niche-Inspired Culture Condition. *Cell Stem Cell* **23,** 787–793.e6 (2018).
4. Fujii, M., Matano, M., Nanki, K. & Sato, T. Efficient genetic engineering of human intestinal organoids using electroporation. *Nature Protocols* **10,** 1474–1485 (2015).
5. Kim, S., Kim, D., Cho, S. W., Kim, J. & Kim, J.-S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* **24,** 1012–1019 (2014).
6. Lin, S., Staahl, B. T., Alla, R. K. & Doudna, J. A. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife* **3,** e04766 (2014).
7. Roberts, B. *et al.* Systematic gene tagging using CRISPR/Cas9 in human stem cells to illuminate cell organization. *Mol. Biol. Cell* **28,** 2854–2874 (2017).
8. Paquet, D. *et al.* Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* **533,** 125–129 (2016).
9. Yao, X. *et al.* Tild-CRISPR Allows for Efficient and Precise Gene Knockin in Mouse and Human Cells. *Dev. Cell* **45,** 526–536.e5 (2018).
10. Yu, C. *et al.* Small Molecules Enhance CRISPR Genome Editing in Pluripotent Stem Cells. *Cell Stem Cell* **16,** 142–147 (2015).
11. Song, J. *et al.* RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat Commun* **7,** 1–7 (2016).
12. Maruyama, T. *et al.* Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* **33,** 538–542 (2015).
13. Li, G. *et al.* Small molecules enhance CRISPR/Cas9-mediated homology-directed genome editing in primary cells. *Sci Rep* **7,** 1–11 (2017).
14. Leonetti, M. D., Sekine, S., Kamiyama, D., Weissman, J. S. & Huang, B. A scalable strategy for high-throughput GFP tagging of endogenous human proteins. *Proc. Natl. Acad. Sci. U.S.A.* **113,** E3501–8 (2016).
15. Kamiyama, D. *et al.* Versatile protein tagging in cells with split fluorescent protein. *Nat Commun* **7,** 1–9 (2016).
16. Dewari, P. S. *et al.* An efficient and scalable pipeline for epitope tagging

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

in mammalian stem cells using Cas9 ribonucleoprotein. *Elife* **7,** 87 (2018).

17. Smith, J. R. *et al.* Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration. *Stem Cells* **26,** 496–504 (2008).
18. Smits, A. H. *et al.* Biological plasticity rescues target activity in CRISPR knock outs. *Nat. Methods* **16,** 1087–1093 (2019).
19. Tuladhar, R. *et al.* CRISPR-Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. *Nat Commun* **10,** 1–10 (2019).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgements

We thank A. Bassett (Wellcome Sanger Institute) for the Cas9 expressing vector. D.S. is supported by a Wellcome Trust PhD studentship (109146/Z/15/Z) and the Department of Pathology, University of Cambridge; L.D.E is supported by the Alzheimer's Research UK Stem Cell Research Centre, funded by the Alborada Trust. E.L.R. is supported by Medical Research Council (MR/P009581/1). Core support from the Wellcome Trust (203144/Z/16/Z) and Cancer Research UK (C6946/A24843).

Contributions

D.S. designed and performed experiments, analysed data, wrote and edited the manuscript. L.D.E. performed experiments. E.L.R conceived and led the project, designed experiments, wrote and edited the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure Legends

Figure 1. CRISPR gene targeting for the *ACTB* locus in human organoids. **(a)** Schematic of the Organoid Easytag workflow. ssRNP and a circular plasmid repair template are nucleofected at day 0. By day 3 cells have proliferated to become tiny colonies and are removed from the Matrigel and dissociated for selection by flow cytometry. $EGFP^{+}$ cells are re-plated sparsely (approx. 1000-1500 cells/well of a 24-well plate) and grown until day 15 when organoids reach a sufficient size to be manually picked under a fluorescent microscope. Organoids are picked into individual wells and passaged until sufficient cells are obtained for both genotyping and freezing down the line. Cells with red nuclei represent incorrectly targeted cells. Cells with white nuclei denote correctly targeted cells. **(b)** Schematic of repair template design for the N terminal fusion *mEGFP-ACTB* gene targeting strategy and the final product. Arrow shows the position of gRNA. E1, exon 1; E2, exon 2; 5'HA, 5' homologous arm; 3'HA, 3' homologous arm. **(c)** Representative flow cytometry results showing the percentage of EGFP cells 72 hrs after nucleofection is performed. **(d)** Representative image showing *mEGFP-ACTB* organoid. DIC channel on the left and EGFP channel on the right. **(e)** Immunofluorescence of *mEGFP-ACTB* organoids. Blue: DAPI (nuclei); green: EGFP (ACTB fusion protein); red: SOX9 (lung progenitor marker). Scale bar denotes 50 µm in all panels.

Figure 2. Generalisation of the Organoid Easytag workflow. **(a)** Schematic of the *SOX9* locus repair template design and final product. E1, Exon 1; E2, Exon 2; E3, Exon 3. 5' HA, 5' homologous arm; 3' HA, 3' homologous arm. Arrow shows the position of the gRNA. **(b)** Immunofluorescence of *SOX9-T2A-H2B-EGFP* organoids. Green: EGFP (SOX9 transcriptional reporter); red: SOX9 (lung progenitor marker); white: ECAD (E-Cadherin, basal-lateral

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

junctions). **(c)** Schematic of the *AAVS1* targeting repair template design and final product. **(d)** Immunofluorescence of *AAVS1-EF1α-mTagRFP-T* organoids. Blue: DAPI (nuclei); red: mTagRFP-T (membrane localised reporter); white: SOX2 (lung progenitor marker). **(e)** Summary of the gene targeting achieved in the human foetal lung organoid system. (**f**) Schematic showing repair template design and final product for generation of *SOX2* knockout organoids using organoid Easytag workflow. Two gRNAs (indicated by arrows) were used at the N and C terminal of *SOX2* CDS respectively. *SOX2* CDS was replaced by *T2A-H2B-EGFP*. **(g)** Immunofluorescence showing that SOX2 protein is completely knocked-out. Blue: DAPI (nuclei); green: EGFP (*SOX2* transcriptional reporter); red: SOX2 (lung progenitor marker); white: ECAD (E-Cadherin, basal-lateral junctions). Scale bar denotes 50 μm (b, d) and 100 μm (g).

Supplementary Figure 1. Optimisation of DNA delivery and CRISPR/Cas9 site specific DNA cleavage efficiency. (a) Efficiency quantification of different DNA delivery methods. EGFP positive cells were quantified by flow cytometry 72 hrs after transfection or transduction. N = 4 different organoid lines were used for each condition. **(b)** Representative images showing DNA delivery efficiencies of different methods 72 hrs after transfection/transduction. Top panel: bright-field channel; bottom panel: GFP channel. Scale bar denotes 100 μm. **(c)** T7 endonuclease assay showing the DNA cleavage efficiency of different forms of Cas9 on the *ACTB* locus. Arrows denote lower bands generated by T7 endonuclease cutting. Left panel: schematic showing the different methods tested for introducing the Cas9 complex. cr/tr RNP, synthetic crispr/tracer RNA heterodimer with Cas9 RNP. ssRNA, single strand synthetic guide RNA with Cas9 RNP. **(d)** Quantification of insertions and deletions (indels) produced by the different CRISPR methods tested using ICE online analysis software (https://ice.synthego.com/). N =3 different organoid lines were used for each condition. Two-tailed Student's t-tests were performed. p

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

values are reported as follows: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.

Supplementary Figure 2. Small molecules did not improve organoid gene targeting efficiency. (a) Workflow for testing different HDR enhancing drugs. Organoid cells were treated with DMSO, RS-1, L755507 and SCR-7 for 48hrs after nucleofection for *mEGFP-ACTB* gene targeting. The percentage of EGFP positive cells was analyzed 72 hrs after nucleofection. DMSO treatment was used as a negative control. (**b**) Summary of EGFP positive percentage. No significant improvement of targeting efficiency was observed after the drug treatments. N =3 different organoid lines were used.

Supplementary Figure 3. ssODN is not suitable for organoid Easytag workflow. (a) Schematic of *GFP11-ACTB* gene targeting using an ssODN repair template. Left panel: the ssRNP, ssODN repair template and a circular plasmid encoding GFP(1-10) were co-transfected. Right panel: the ssRNP, plasmid repair template and a circular plasmid encoding GFP(1-10) were co-transfected. **(a')** Schematic of split GFP system. In the correctly targeted cells, if GFP(1-10) is also present, GFP11-ACTB and GFP(1-10) combine to form a fluorescent protein. However, if either component is missing, cells do not show fluorescence. (**b**) Representative FACS results showing *GFP11-ACTB* gene targeting using different repair templates. In the absence of repair template, GFP(1-10) alone presents a minimal background level of GFP fluorescence (top-right panel). In both the ssODN group, or the plasmid group, clear GFP^+ cells can be seen in the flow cytometry plots. (**c**) Representative bright-field images showing the organoid colonies forming 10 days after GFP^+ cells were put into culture. In the ssODN group, no obvious colonies were observed. Whereas, in the plasmid group, numerous small colonies started to become visible. White arrowheads indicate organoid colonies. Scale bar denotes 2 mm**.**

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(**d**) Schematic drawing of *SOX2-V5* repair template ssODN design. The V5 tag is designed to be inserted at the C-terminal of the SOX2 protein. A 'beacon vector' encoding GFP is co-transfected together with the ssODN to enrich for SOX2-V5 correctly targeted cells. (**e**) Genotyping gel results to show Clone #3 is a potential heterozygote for *SOX2-V5* gene targeting. (**e'**) Illustration of sequencing results for Clone #3. A random insertion is found before the V5 tag sequence (bottom) compared with the expected correct V5 tag targeting (top).

Supplementary Figure 4. Characterisation of *SOX9* targeted colonies. (a) Immunofluorescence of *SOX9-T2A-H2B-EGFP* clone. Blue: DAPI (nuclei); green: H2B-EGFP (*SOX9* transcriptional reporter); red: NKX2-1 (lung epithelial marker); white: SOX2 (lung progenitor marker). Scale bar denotes 50 μm. **(b)** Schematic of the *SOX9* genotyping strategy. 5' and 3' junction amplicons consist of a primer inside the EGFP sequence and another primer upstream, or downstream, of the homologous arm respectively. **(c)** Representative gel image showing correct colony genotyping results. **(d)** Illustration of Sanger sequencing results for *SOX9-T2A-H2B-EGFP* heterozygous WT allele, focusing on the gRNA cutting sites. A single point mutation outside the gRNA sequence in the 3'UTR of SOX9 is observed in 4/6 of the clones sequenced #3-6). Results were from 6 lines, randomly picked from 3 different parental organoid lines.

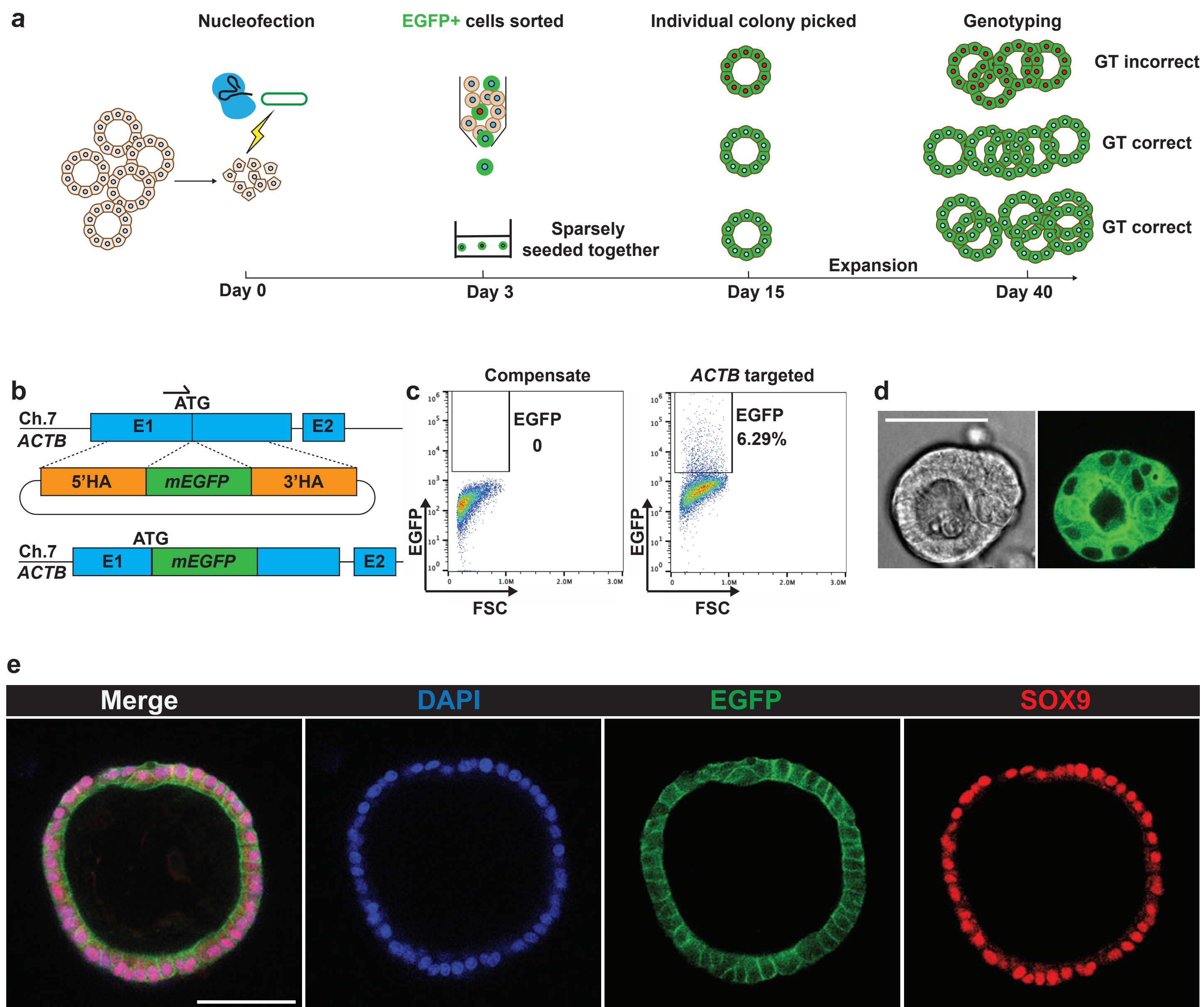
Supplementary Figure 5. Generation of *SOX2* knockout using organoid Easytag workflow. (a) Genotyping strategy for *SOX2 +/H2B-EGFP* heterozygotes. **(b)** Representative gel image showing correctly targeted colonies. **(c)** Representative flow cytometry results showing the EGFP signal shift from WT control to *SOX2 +/H2BEGFP* Clone #5 heterozygote and *SOX2 +/H2BEGFP* Clone #5 retargeted to remove the second copy of *SOX2 CDS*. **(d)** Genotyping strategy for testing whether the WT *SOX2 CDS* is still present. **(e)**

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Representative gel image showing correctly targeted colonies (Clone #5 and # 6) with no amplification for WT *SOX2 CDS* in contrast to WT control.

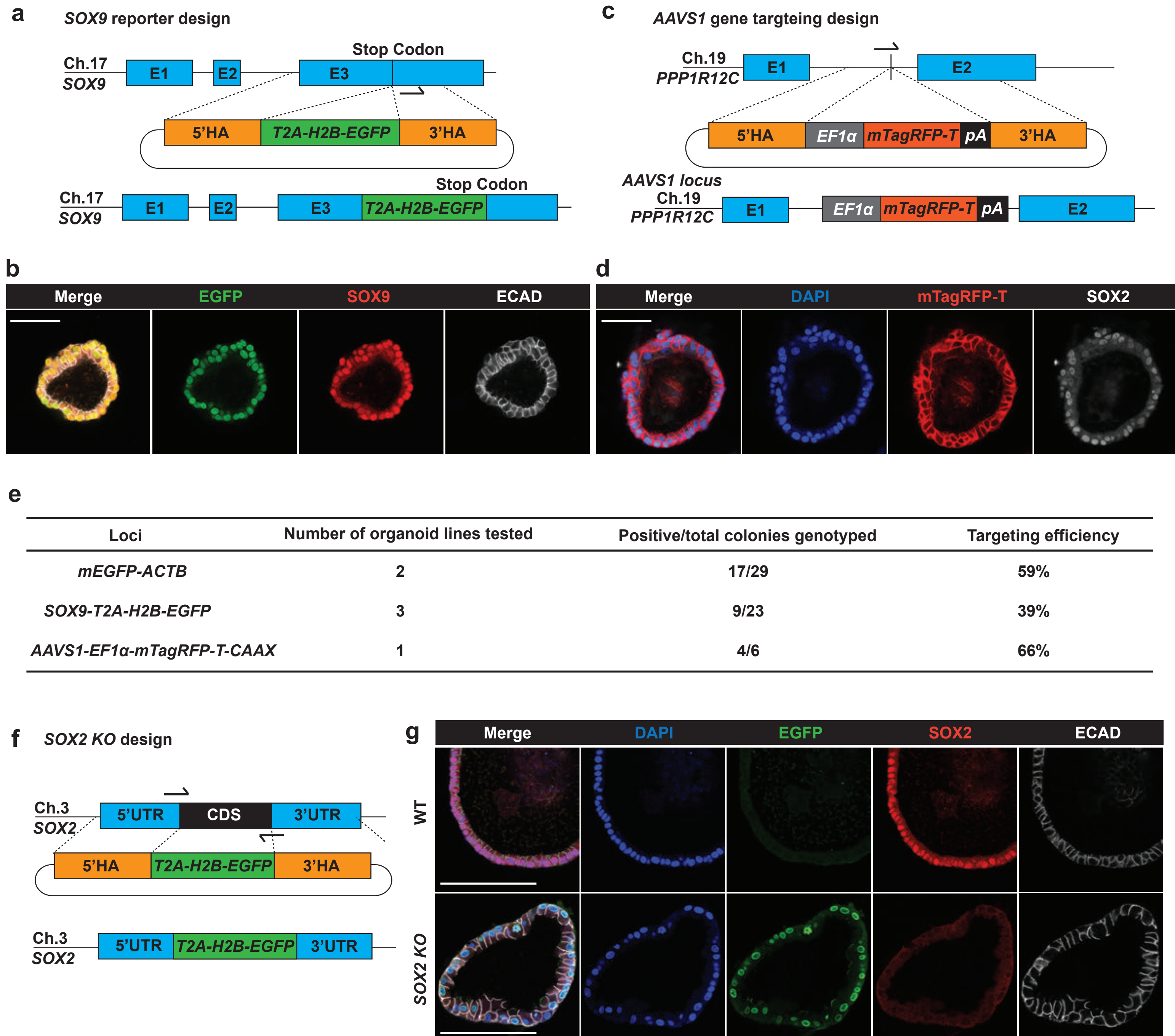
bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1. CRISPR gene targeting for *ACTB* locus in human organoids.



bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

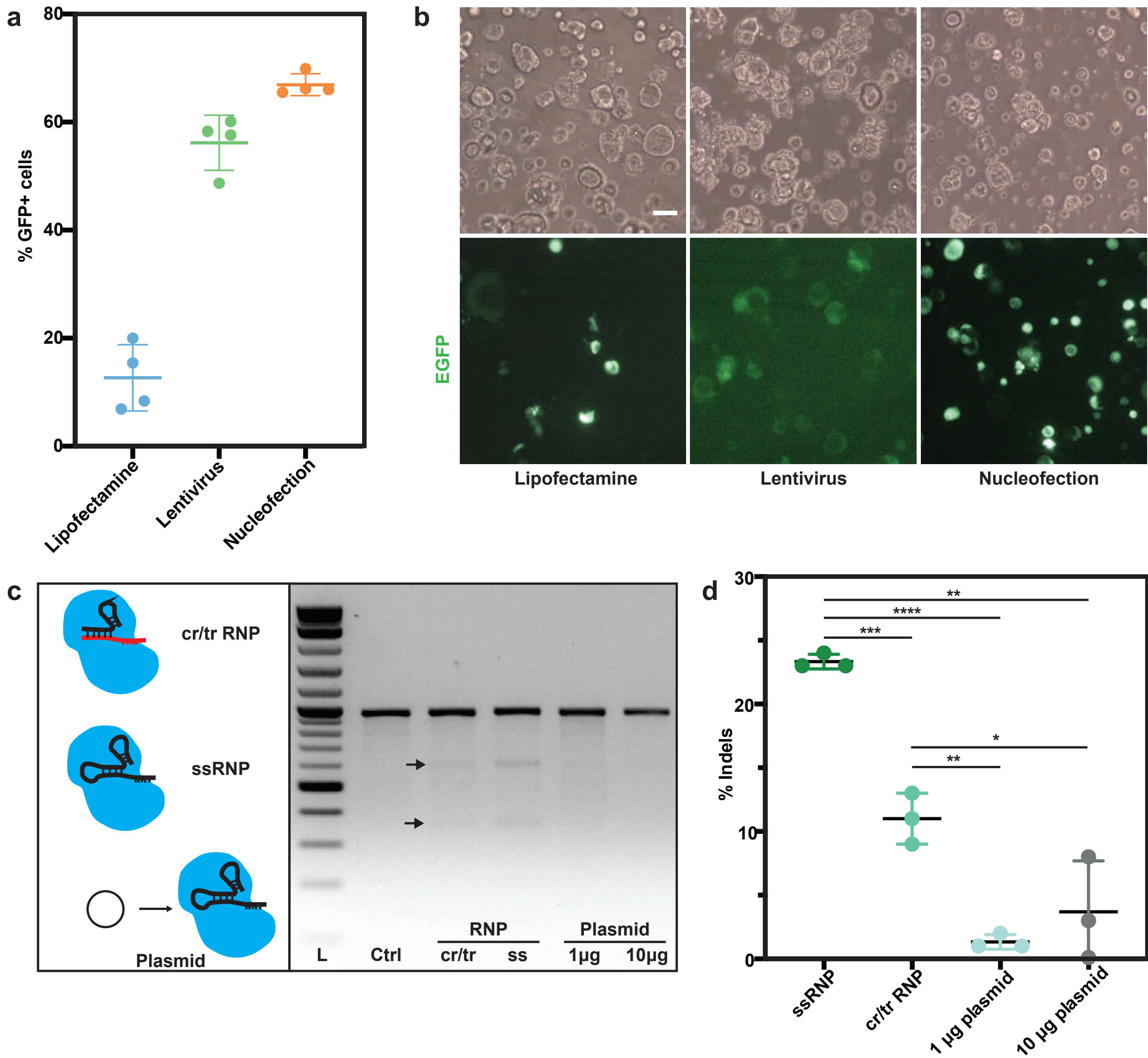
Figure 2. Generalisation of Organoid Easytag workflow.



Loci	Number of organoid lines tested	Positive/total colonies genotyped	Targeting efficiency
mEGFP-ACTB	2	17/29	59%
SOX9-T2A-H2B-EGFP	3	9/23	39%
AAVS1-EF1α-mTagRFP-T-CAAX	1	4/6	66%

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

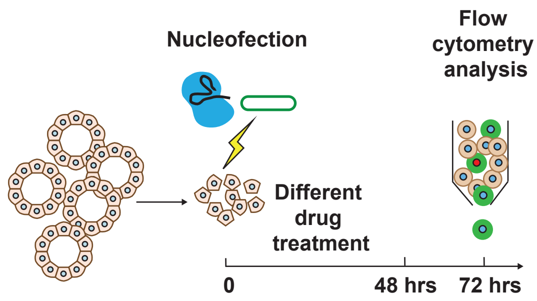
Supplementary Figure 1. Optimisation of DNA delivery and CRISPR/Cas9 site specific DNA cleavage efficiency.



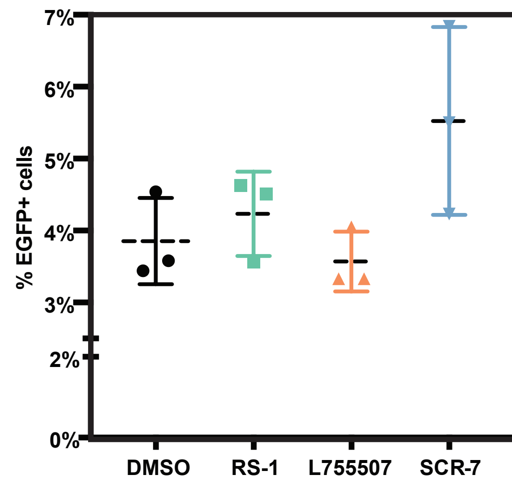
bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Figure 2. Small molecules did not improve organoid gene targeting efficiency.

a

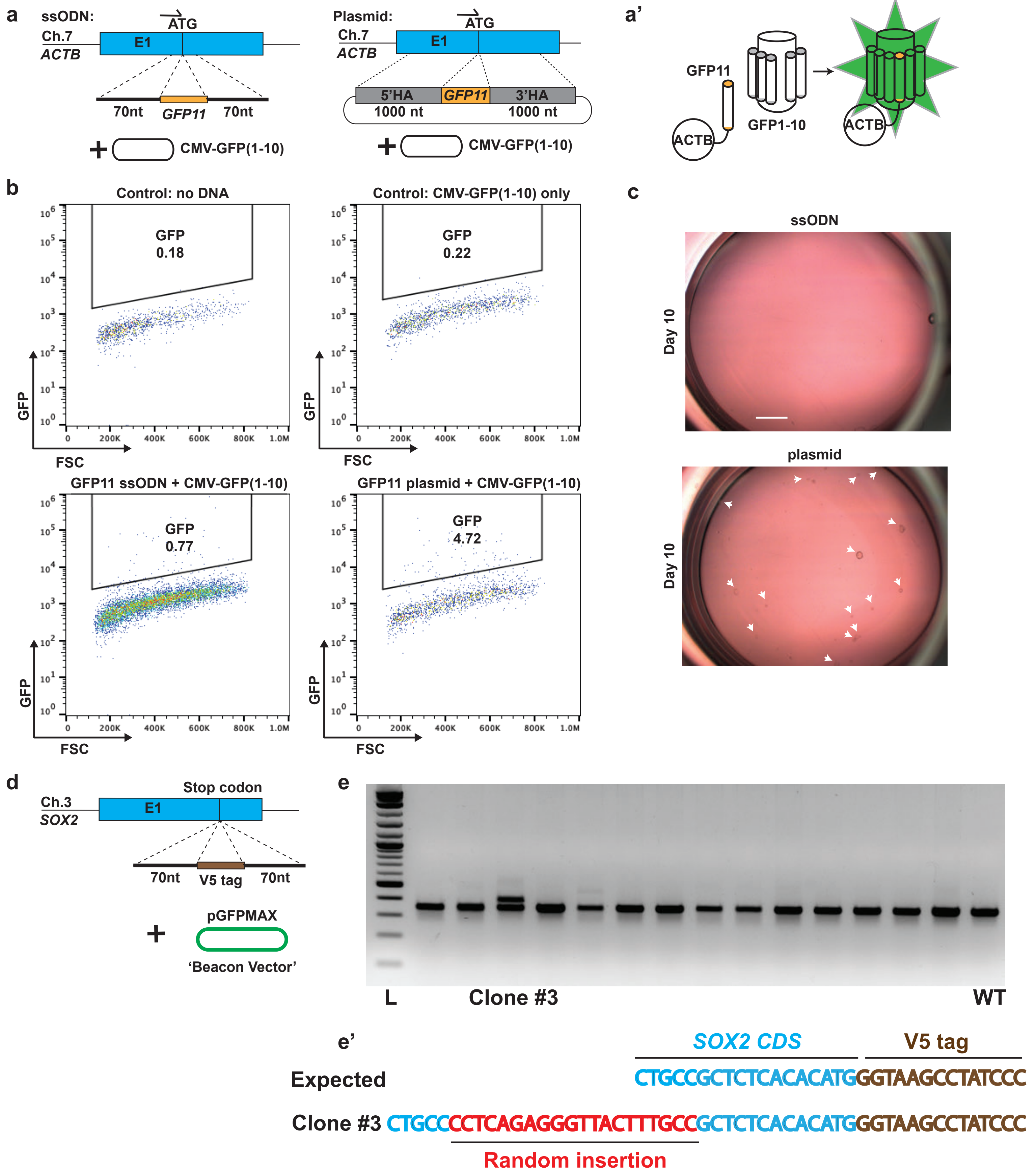


b



bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

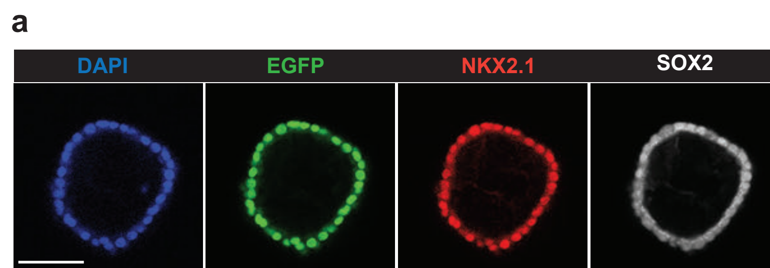
Supplementary Figure 3. ssODN is not suitable for organoid Easytag workflow.



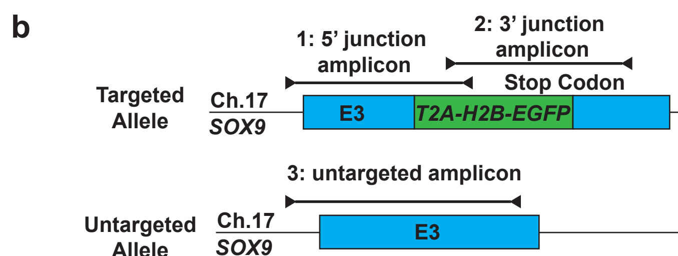
bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Figure 4. Characterisation of SOX9 targeted colonies.

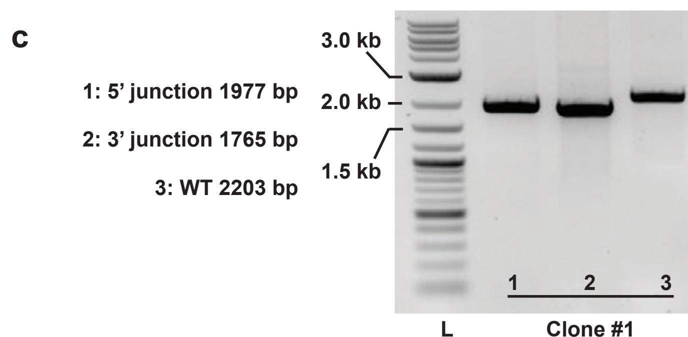
a



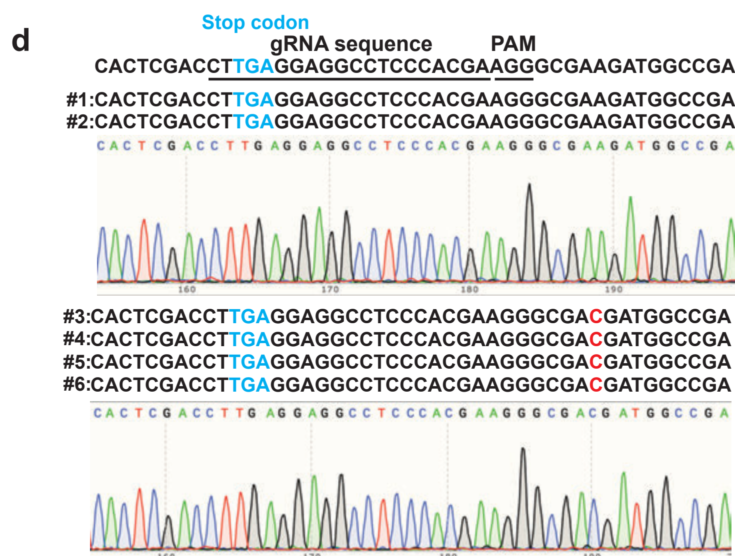
b



c



d



bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.076067; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Figure 5. Generation of SOX2 knockout using organoid Easytag workflow.