FEBS Letters

FEBS PRESS science publishing by scientists

Simultaneous generation of multi-gene knockouts in human cells

Yuexin Zhou[1,†], Hongmin Zhang[1,2,†] and Wensheng Wei[1]

1 Biodynamic Optical Imaging Center (BIOPIC), Beijing Advanced Innovation Center for Genomics, Peking-Tsinghua Center for Life Sciences, State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing, China
2 Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China

Correspondence
W. Wei, Biodynamic Optical Imaging Center (BIOPIC), Beijing Advanced Innovation Center for Genomics, Peking-Tsinghua Center for Life Sciences, State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing 100871, China
Fax: +86 10 62757131
Tel: +86 10 62757227
E-mail: wswei@pku.edu.cn

[†]These authors contributed equally to this work.

(Received 7 October 2016, revised 16 October 2016, accepted 18 October 2016)

doi:10.1002/1873-3468.12469

Edited by Ned Mantei

Genome-editing techniques enable the generation of gene knockouts in various mammalian cell lines. However, it remains technically challenging to completely disrupt a targeted gene using a canonical method in a timely manner. To improve the efficiency of producing reliable genomic modifications, we designed a method using a linear donor fragment containing a reporter system. Combined with a homologous recombination-independent knock-in strategy, we successfully enriched those cell clones that specifically carry the target gene mutations. We observed a much improved success rate when generating single- and multiple-gene knockouts in a one-step procedure using this special protocol coupled with the CRISPR/Cas9 system. This new approach further empowers the molecular biological study of genes and their functions.

Keywords: CRISPR; knockout; nonhomologous end-joining

Genome-editing techniques have revolutionized the experimental interrogation of gene function. The three main techniques, ZFN (zinc-finger nucleases) [1], TALENs (transcription activator-like effector nucleases) [2–4] and the CRISPR/Cas system [5–7], employ different mechanisms to generate sequence-specific double-strand breaks (DSBs) and subsequently trigger native repair systems to accomplish sequence-specific modifications [8,9]. These powerful technologies have broad applications in functional gene studies [10], dynamic imaging of chromosome loci in live cells [11,12], correction of disease mutation [13], gene therapies [14] and beyond. The CRISPR/Cas system has become particularly popular because of its high efficiency and ease of operation. However, although the CRISPR/Cas system has unprecedented power in design-based and sequence-specific genome interrogation, it remains technically challenging even for the simple task of creating gene knockouts in mammalian cells [15]. Various efforts have been made to improve the efficiency of protocols to generate gene knockouts, such as integration of the CRISPR/Cas system for persistent expression of Cas9 and sgRNA [16], pregeneration of a Cas9 stable-expression cell line [17], enhancement of the nonhomologous end-joining (NHEJ) pathway [18], enrichment of the gene-target

Abbreviations
DSBs, double-strand breaks; DT, diphtheria toxin; FACS, fluorescence-activated cell sorting; HR, homologous recombination; NHEJ, nonhomologous end-joining; pgRNA, paired guide RNA; sgRNA, single guide RNA; TALENs, transcription activator-like effector nucleases; ZFN, zinc-finger nucleases.

event through a simultaneous disruption of a separate gene that enables a specific drug selection [19] and the usage of a surrogate reporter to enrich the gene knockout [20,21]. However, various shortcomings limit the general and broad applications of these techniques, and even knocking out a single gene can sometimes be a long, tedious and high-risk process [22]. In particular, it remains a daunting task, if not impossible, to generate multi-gene knockouts in mammalian cells.

If the target gene disruption confers a phenotypic change that can be used for enrichment, gene knockout clones are readily obtained, such as HeLa *CSPG4*$^{-/-}$ cells conferring resistance to *Clostridium difficile* toxin B [23]. However, this strategy cannot be generally applied. The conventional method to obtain gene knockout is to cotransfect Cas9- and sgRNA-expressing plasmids with a plasmid expressing antibiotic resistance or fluorescent proteins [23,24], but this selection would not increase the mutation rate or enrich those small portions of cells containing target modifications. It has been previously reported that the mutation frequency of target alleles is generally high if one of the homologous alleles is modified [25,26]. We reasoned that if we could insert a donor in the specific site on one of the target alleles and select clones expressing a donor-inherited marker gene, we might be able to enrich those rare events in which all alleles are modified. It has been shown that exogenous dsDNA fragments can be integrated into chromosomal loci with DSBs through different repair mechanisms. The integration efficiency is generally much higher through NHEJ-mediated DNA repair [27,28] than through homologous recombination (HR) repair [29]. To ensure the efficiency for donor insertion and to avoid the hassle of cloning long flanking sequences for donor construction, we used the CRISPR/Cas9-triggered HR-independent DNA repair to mediate the insertion of exogenous linear donor DNAs.

Methods

Cell culture and transfection

HeLa, $HeLa_{OC}$ and HEK293T cells were maintained in Dulbecco's modified Eagle's medium (10-013-CV; Corning, Tewksbury, MA, USA) supplemented with 10% foetal bovine serum (Lanzhou Bailing Biotechnology Co., Ltd., Lanzhou, China) with 5% CO_2 at 37 °C. For transfection, all cells were seeded on six-well plates and transfected with X-tremeGENE HP (06366546001; Roche, Mannheim, German) according to the supplier's protocols. Briefly, 2 μg of DNA and 4 μL of X-tremeGENE HP were added to 200 μL of Opti-MEM® I Reduced Serum Medium (31985088; Thermo Fisher Scientific, Grand Island, NE, USA). The mixture was incubated at room temperature for 15 min and then added to the cells.

gRNA-expressing plasmid cloning

For sgRNA-expressing plasmids, the oligonucleotides for each sgRNA-coding sequence were individually designed (Table S2) and synthesized (Ruibiotech, Inc., Beijing, China). We dissolved the oligonucleotides with 1 × TE to a concentration of 10 μM, mixed paired oligonucleotides with TransTaq HiFi Buffer II (K10222; TransGen Biotech, Beijing, China), heated to 95 °C for 3 min and then slowly cooled to 4 °C. We phosphorylated these annealed oligonucleotide pairs at 37 °C for 30 min, and after heat inactivation, ligated the product into the sgRNA vector using the Golden Gate method. For pgRNA-expressing plasmids, primers containing two gRNA-coding sequences were used to amplify the scaffold sequence of the gRNA and the U6 promoter (Table S2). The PCR products were then purified and ligated into the backbone using the Golden Gate method. Compared to the sgRNA backbone vector we reported before [17], the sgRNA scaffold of the new vector was modified [30], and the EGFP-coding sequence was replaced with an mCherry-coding sequence.

T7E1 assay

Genomic DNA was extracted using DNeasy Blood & Tissue Kit (69504; Qiagen, Hilden, Germany) for PCR amplification of the genome region containing the gRNA-targeting sequence. The primer sequences used in these assays are listed in Table S1.

Linear donor construction

Donors containing the CMV-driven puromycin-resistant gene or the EGFP gene and sequences carrying stop codons in three reading frames were pregenerated and cloned into the pEASY-T5-Zero cloning vector (CT501-02; TransGen Biotech), serving as a universal template (Appendix S1). Primers containing sgRNA cutting site(s) and protection sequences were used to amplify the template by two-step PCR reactions. The primer sequences are listed in Table S4. We performed the PCR reaction using the *Trans Taq*® DNA Polymerase High Fidelity (HiFi) kit (K10222; TransGen Biotech). After PCR reaction, the donors were purified using the DNA Clean & Concentrator-25 kit (D4034; Zymo Research, Orange, CA, USA).

HR-independent donor insertion and cell selection

For $HeLa_{OC}$ cells, 1 μg of purified linear donor PCR product and 1 μg of sgRNA/pgRNA were transfected into the cells,

and cells were treated with 1 μg·mL^{-1} puromycin 2 weeks after transfection. For HeLa and HEK293T cells, 1 μg of donor and 0.5 μg of sgRNA/pgRNA with 0.5 μg of Cas9 plasmid were transfected into the cells. Cells were then treated 2 weeks after transfection with 1 μg·mL^{-1} of puromycin or fluorescence-activated cell sorting (FACS) to determine EGFP-positivity depending on the type of donor.

Conventional gene knockout protocols

For HeLa cells, Cas9 plasmid, sgRNA and a puromycin-resistant plasmid (0.9 μg: 0.9 μg: 0.2 μg) were transfected into the cells, and cells were selected with 1 μg·mL^{-1} puromycin 2 days after transfection. For HeLa$_{OC}$ cells, sgRNA and a puromycin-resistant plasmid (1.8 μg: 0.2 μg) were transfected into the cells, and cells were selected with 1 μg·mL^{-1} puromycin 2 days after transfection. For HEK293T cells, Cas9 plasmid, sgRNA and a plasmid expressing mCherry (0.9 μg: 0.9 μg: 0.2 μg) were transfected into the cells, and cells were selected with FACS 2 days after transfection.

Splinkerette PCR

Genomic DNA from cells was extracted using DNeasy Blood & Tissue Kit (69504; Qiagen). Two microgram of genomic DNA was digested by *Mlu*I and *Nsi*I (R0198V and R0127V; NEB, Beverly, MA, USA) at 37 °C for 6 h, followed by 65 °C inactivation for 20 min. We added the splinkerette adaptor and T4 ligase (M0202L; NEB) to the digested products, which were purified using DNA Clean & Concentrator-25 kit (D4034; Zymo Research). After incubation at 16 °C overnight, we performed the primary PCR using primers splinker1/R1 on the purified ligation products. After purification, primary PCR product was used as template for the secondary PCR reaction using primers splinker2/R1. More detailed methods have been previously described [31–33]. Sequences of the primers and adaptor are listed in Table S5.

Quantitative real-time PCR (qRT-PCR)

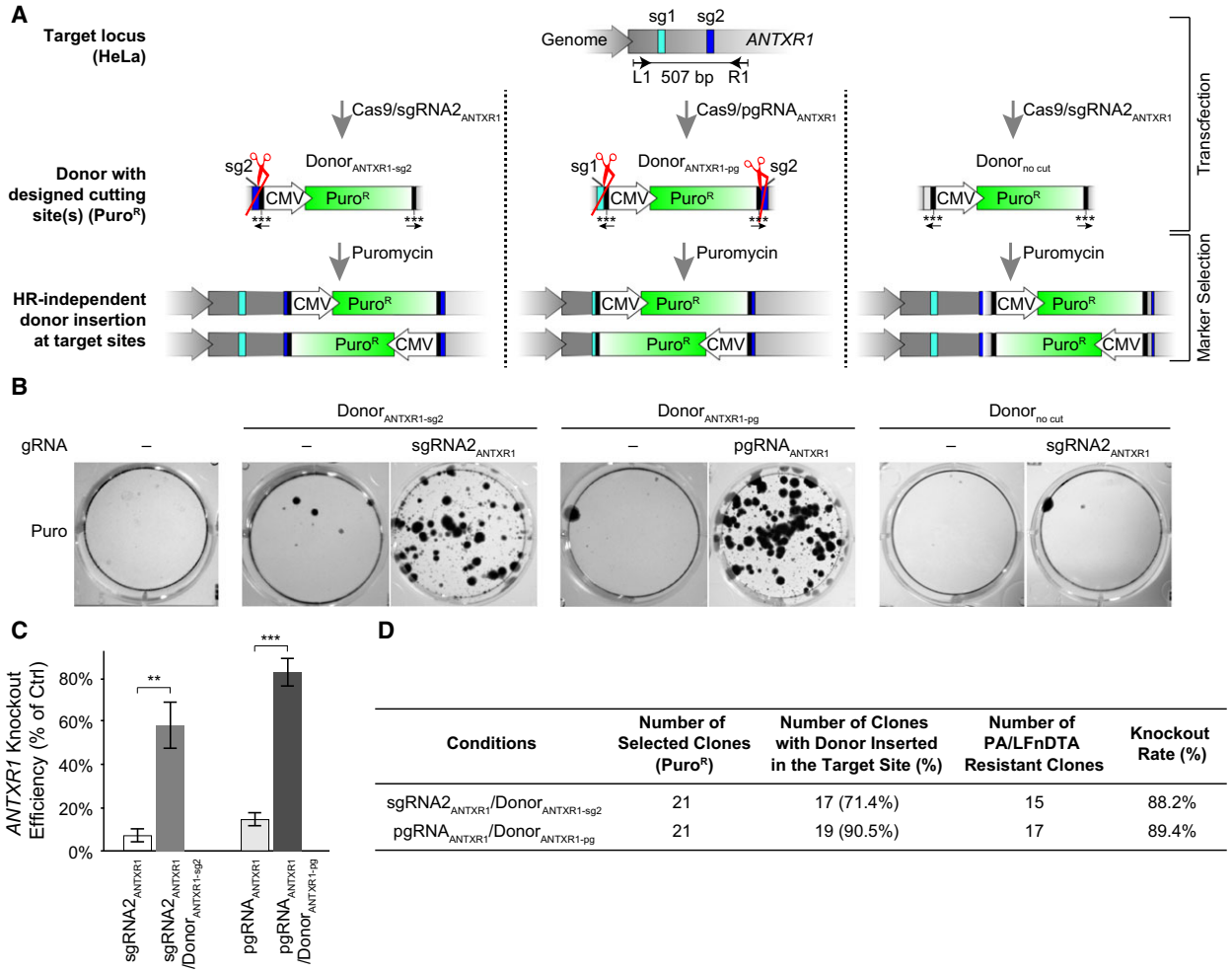
RNA of cultured cells was extracted using RNAprep Pure Micro kit (DP420; TIANGEN, Beijing, China), and the cDNA was synthesized using QuantScript RT kit (KR103-03; TIANGEN). SYBR Premix Ex Taq II (RR820A; TaKaRa, Otsu, Japan) were used to perform real-time PCR on LightCycler96 qPCR system (Roche). Transcript levels of GAPDH were measured as normalized controls. Specific primers (5′-ATTGAGGAGGTAGATTAGGGGC-3′/5′-CACAGGAAATTGAGAACTGACAAAC-3′ for HSPA1A and 5′-CATTGAGGAGGTGGATTAGGGG-3′/5′-AAGAAGTGAAGCAGCAAAGAGC-3′ for HSPA1B) were used for qRT-PCR.

Western blot analysis

For western blot analysis, cell lysates were obtained, resolved on SDS/PAGE gels (1610183; Bio-Rad, Hercules, CA, USA) and transferred onto a polyvinylidene fluoride membrane (IPVH00010; Millipore, Billerica, MA, USA) as previously described. Membranes were blocked with 5% nonfat milk at 37 °C for 1 h, probed with primary anti-HSPA1L antibody (ab154409; Abcam, Cambridge, MA, USA) and anti-β-tubulin antibody (CW0098M; CWBIO, Beijing, China) overnight at 4 °C, and then individually incubated with goat anti-rabbit IgG-HRP secondary antibody (111035003; Jackson Immunoresearch, West Grove, PA, USA) and goat anti-mouse IgG-HRP secondary antibody (115035003; Jackson Immunoresearch) at room temperature for 1 h. Protein bands were then detected using a Clarity™ Western ECL Substrate kit (1705060; Bio-Rad) and imaged with the Chemi-doc system (1708370; Bio-Rad).

Results

We planned to generate linear donor DNA and to test the effect of site-specific DNA integration for the potential enrichment of knockout events on the *ANTXR1* gene in HeLa cells. We designed two sgRNAs targeting the first exon of *ANTXR1* and verified their efficiency in creating indels on target loci (Table S1). The linear donor consists of four parts as follows (from the inside out): a puromycin-resistant gene driven by a CMV promoter in the middle, the short stretch of sequence (at both sides) consisting of stop codons in all three reading frames, a single guide RNA (sgRNA) recognition sequence at the 5′ end or paired guide RNA (pgRNA) recognition sequences at both sides, and 20-bp protection sequences at both ends. We expected that this linear donor (Donor$_{ANTXR1\text{-}sg2}$ or Donor$_{ANTXR1\text{-}pg}$) (Fig. 1A) could be integrated into the targeting site in either orientation. As a control, we purposely engineered a donor (Donor$_{no\ cut}$) whose sgRNA recognition sequence was replaced by a random 20-bp sequence that could not be recognized by either sgRNA (Fig. 1A). After cotransfection of the Cas9-expressing plasmid, sgRNA2$_{ANTXR1}$ or pgRNA$_{ANTXR1}$ and their corresponding donors, Donor$_{ANTXR1\text{-}sg2}$ or Donor$_{ANTXR1\text{-}pg}$, we cultured the cells for 14 days before applying antibiotics treatment for 2 days, so that we could select clones with donor integration other than those that have either puro^{+} transient expression or adapted antibiotic resistance. We then counted the numbers of puromycin-resistant (puro^{+}) clones that were stained by MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide). Transfection with donor alone without sgRNA/pgRNA produced very few puromycin-resistant clones,



Conditions	Number of Selected Clones (Puro^R)	Number of Clones with Donor Inserted in the Target Site (%)	Number of PA/LFnDTA Resistant Clones	Knockout Rate (%)
$sgRNA2_{ANTXR1}$/$Donor_{ANTXR1\text{-}sg2}$	21	17 (71.4%)	15	88.2%
$pgRNA_{ANTXR1}$/$Donor_{ANTXR1\text{-}pg}$	21	19 (90.5%)	17	89.4%

Fig. 1. Donor design and application for puromycin selection-mediated enrichment of cells containing a Cas9/gRNA-targeted mutation at the *ANTXR1* locus in HeLa cells. (A) Schematic of the homologous recombination (HR)-independent knock-in of a linear donor on the sgRNA- or pgRNA-targeting site of the *ANTXR1* gene. The guide RNA cutting sites are labelled on both the genome locus and the linear donors as indicated. The short stretches of sequences with stop codons in all three reading frames (3 × STOP) are labelled with ***, and the arrows point in the direction of the reading frames. The L1/R1 primer sequences used for PCR amplification are listed in Table S3. (B) MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) staining of puromycin-resistant colonies for cells transfected with sgRNA/pgRNA with or without its donor as indicated. (C) *ANTXR1* knockout rate as a percentage of cells resistant to PA/LFnDTA. HeLa cells were transfected as indicated with sgRNA/pgRNA with corresponding donors (dark bar) or with a puromycin-resistant plasmid (light bar). Cells were selected with puromycin (1 $\mu g \cdot mL^{-1}$) prior to the PA/LFnDTA resistance analysis. Error bars indicate SD ($n = 3$), *t*-test, $^{**}P < 0.01$, $^{***}P < 0.001$. (D) Summary of *ANTXR1* knockout cell enrichment using different gRNAs and their donors.

likely due to rare and random integration of the linear donor into chromosomes; however, the number of these clones was much lower than in cells cotransfected with Cas9, sgRNA/pgRNA and its corresponding donor. In addition, the control donor failed to produce a significant amount of puro$^+$ clones, suggesting that the sgRNA-directed Cas9 cleavage on the donor is critical for effective donor integration (Fig. 1B). $pgRNA_{ANTXR1}$-mediated dual-cutting coupled with $Donor_{ANTXR1\text{-}pg}$ integration appeared to have similar efficiency with $sgRNA2_{ANTXR1}$ plus $Donor_{ANTXR1\text{-}sg2}$, either approach yielded enough puro$^+$ clones for subsequent mutant identification.

Because the *ANTXR1* gene knockout in HeLa cells results in cell resistance to chimeric anthrax toxin, PA/LFnDTA [17], we were able to quickly determine the *ANTXR1* knockout efficiency by calculating the percentage of toxin-resistant cells in the puro$^+$ population. The use of donors with new protocol increased the gene knockout efficiency six- to eightfold compared to conventional gene knockout protocols without donor-mediated enrichment (transfection with

sgRNA2$_{ANTXR1}$ or pgRNA$_{ANTXR1}$ alone) (Figs 1C and S1A). The majority of clones isolated individually from the puro^{+} cell pool contained the donor insertion at the sgRNA-targeting site (Figs 1D and S1B), and close to 90% of cells carrying the donor fragment were genuine gene knockout clones (Fig. 1D).

Because donors with either single- or dual-cutting sites were both able to greatly enhance the selection of cells with modifications at the target loci, we used only the single-cutting donor for convenience. We tested the same protocol by targeting another gene, *HBEGF* (Fig. S2A), and observed the same effect, i.e., that only sgRNA plus its corresponding donor produced abundant puro^{+} clones. This result again demonstrated that donor insertion depends on specific sgRNA/Cas9-mediated DSBs on both targeted loci and the linear donor (Fig. S2B). Because the *HBEGF* gene encodes the diphtheria toxin (DT) receptor and its knockout in HeLa cells results in cell resistance to DT [17], we were able to quickly calculate the knockout rate. Again, when coupled with its corresponding donor, sgRNA greatly improved the *HBEGF* knockout rate (Fig. S2C, D). This new approach is not limited to specific cell types or marker genes; we also witnessed much improved performance when we targeted the *HBEGF* gene in HEK293T cells with donors containing *EGFP* instead of the puromycin-resistant gene (Fig. S3).

We repeated the same protocol, targeting genes using the HeLa$_{OC}$ line that was pregenerated for the constant expression of Cas9 [17]. We obtained similar results as with HeLa cells for the knockout of either *ANTXR1* (Fig. S4) or *HBEGF* (Fig. S5). This method offers a convenient alternative for creating target knockouts. Notably, most of the clones only contained one donor insertion. However, the vast majority of alleles were edited at the targeting sites from donor-positive clones, whereas donor-free protocols appeared to be much less efficient at creating indels (Table 1). This finding clearly

Table 1. Summary of the genomic sequencing results of the *ANTXR1* guide RNA target region in HeLa$_{OC}$ cells.

Category	Sequencing results of PCR fragments with sizes similar to wild-type (~ 500 bp) Wild-type | Mutant | Mutation rate	Sequencing results of PCR fragments with the sizes similar to wild-type plus donor insert (~ 1.8 kb) Wild-type | Mutant | Mutation rate
sgRNA1$_{ANTXR1}$	15 | 0 | 0%	– | – | –
sgRNA1$_{ANTXR1}$/ Donor$_{ANTXR1\text{-}sg1}$	1 | 14 | 93%	15 | 15 | 100%
sgRNA2$_{ANTXR1}$	13 | 2 | 15%	– | – | –
sgRNA2$_{ANTXR1}$/ Donor$_{ANTXR1\text{-}sg2}$	5 | 10 | 67%	15 | 15 | 100%
pgRNA$_{ANTXR1}$	10 | 5 | 33%	– | – | –
pgRNA$_{ANTXR1}$/ Donor$_{ANTXR1\text{-}pg}$	1 | 14 | 93%	15 | 15 | 100%

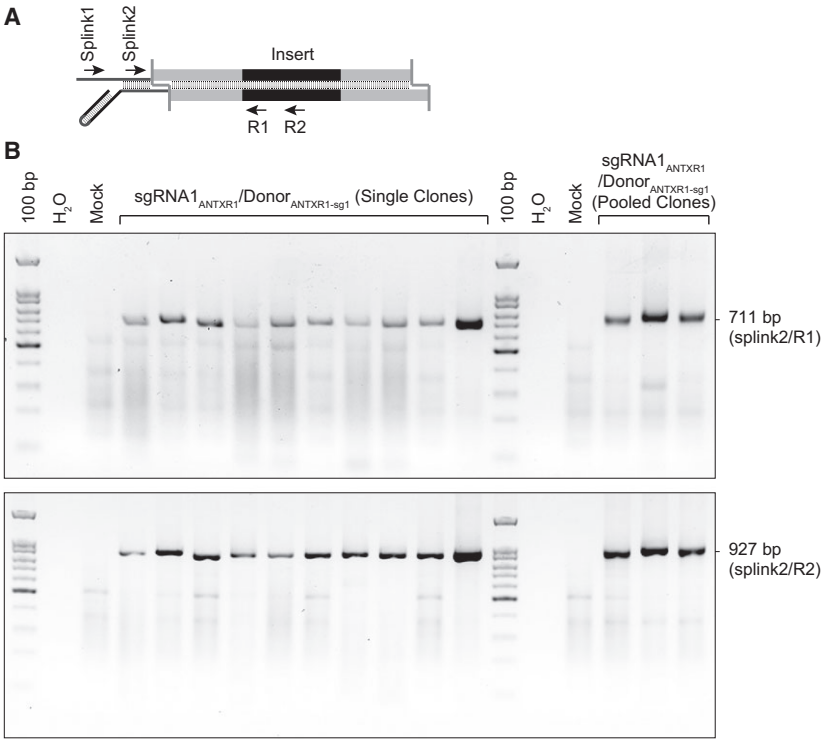


Fig. 2. Off-target assessment of donor insertion using splinkerette PCR (spPCR) analysis in HeLa cells. (A) Design of the adaptor and primers used for spPCR analysis. The Splink1 and Splink2 primers matched the adaptor sequences, and the primers R1 and R2 matched the linear donor sequences. (B) spPCR reactions for the verification of off-target insertion of the linear donor in single clones and pooled cell clones after puromycin selection. The insertion of the donor at the targeted *ANTXR1* locus generated 711- and 927-bp PCR products using the primers Splink2/R1 and Splink2/R2 respectively.

demonstrated that sgRNA or pgRNA action to generate indels tightly correlates with the occurrence of donor insertion, making it an ideal reporter for the enrichment of rare mutant events.

Analysis of junction sequences revealed that most PCR-amplified donor fragments at the integrated sites had a variety of indels at both ends, whereas the targeted genome sites had no or very few changes. The higher rate of indels in donors compared to the targeted chromosomal loci is likely because the linear donors are more vulnerable for degradation by exonucleases. In addition, the donor appeared in both orientations after insertion (Fig. S6), indicating that DSBs from donors, but not the sgRNA recognition sequences, are responsible for donor integration and target gene mutations. Therefore, the donor integration is unlikely through the HR or MMEJ pathway [29,34,35].

To examine whether the use of an external donor influences the off-target effects of the CRISPR/Cas system, we performed splinkerette PCR analysis to survey the genome-wide integration site [31–33]. Correct donor insertion at the *ANTXR1* locus would produce 711- and 927-bp products using the primers Splink2/R1 and Splink2/R2 respectively (Fig. 2A). To perform the splinkerette PCR analysis, we randomly selected 10 single clones with donor insertions and three puro$^+$ pooled clones in HeLa$_{OC}$ cells targeting *ANTXR1* (Fig. S4). Similar to those without donor transfection, single clones or pooled populations through donor enrichment had no detectable off-target effects based on the splinkerette PCR results (Fig. 2B).

To test the efficacy of our new protocol in producing double-gene knockouts, we simultaneously targeted two genes, *PSEN1* and *PSEN2*, in HeLa$_{OC}$ cells. Two sgRNAs targeting each gene were

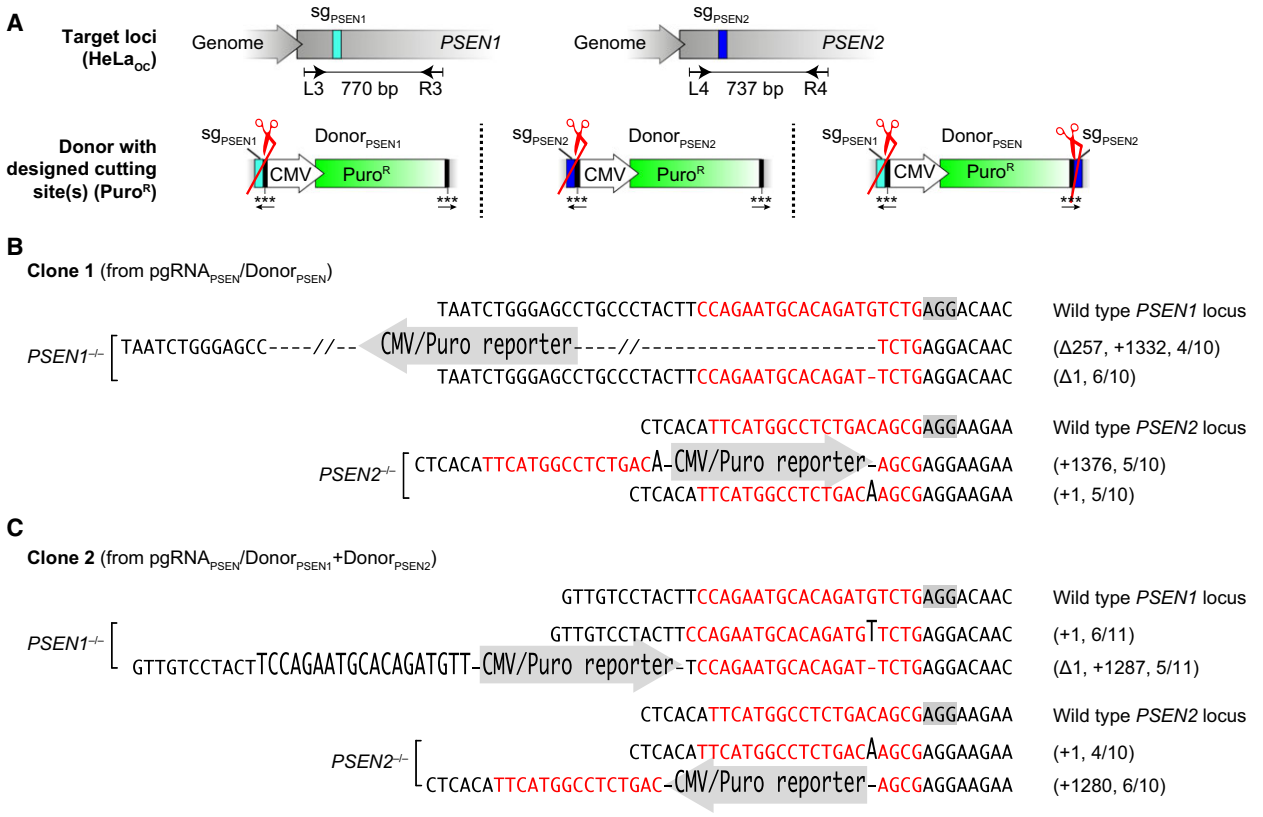


Fig. 3. One-step generation of *PSEN1* and *PSEN2* knockouts in HeLa cells. (A) Schematic of the homologous recombination (HR)-independent knock-in of linear donor(s) on the sgRNA- or pgRNA-targeting sites of *PSEN1* and *PSEN2* genes. Donor$_{PSEN1}$ or Donor$_{PSEN2}$ has the sgRNA$_{PSEN1}$ or sgRNA$_{PSEN2}$ cutting site at its 5′ end. Donor$_{PSEN}$ has the sgRNA$_{PSEN1}$ cutting site at the 5′ end and the sgRNA$_{PSEN2}$ cutting site at the 3′ end. (B and C) Partial coding sequences of the *PSEN1* and *PSEN2* genes in the genome containing the sgRNA-coding regions (red) and sequencing analysis of the mutated alleles. Clone 1 (B) was from HeLa$_{OC}$ cells transfected with pgRNA$_{PSEN1+PSEN2}$/Donor$_{PSEN}$, and clone 2 (C) was from HeLa$_{OC}$ cells transfected with pgRNA$_{PSEN1+PSEN2}$/Donor$_{PSEN1}$+Donor$_{PSEN2}$. Shaded nucleotides represent the PAM sequence that guided the Cas9 for DNA recognition and cleavage. The dashes indicate deletions, and the tall letters indicate nucleotide insertions. Light-grey arrows in the background indicate the direction of the CMV promoter in the donor.

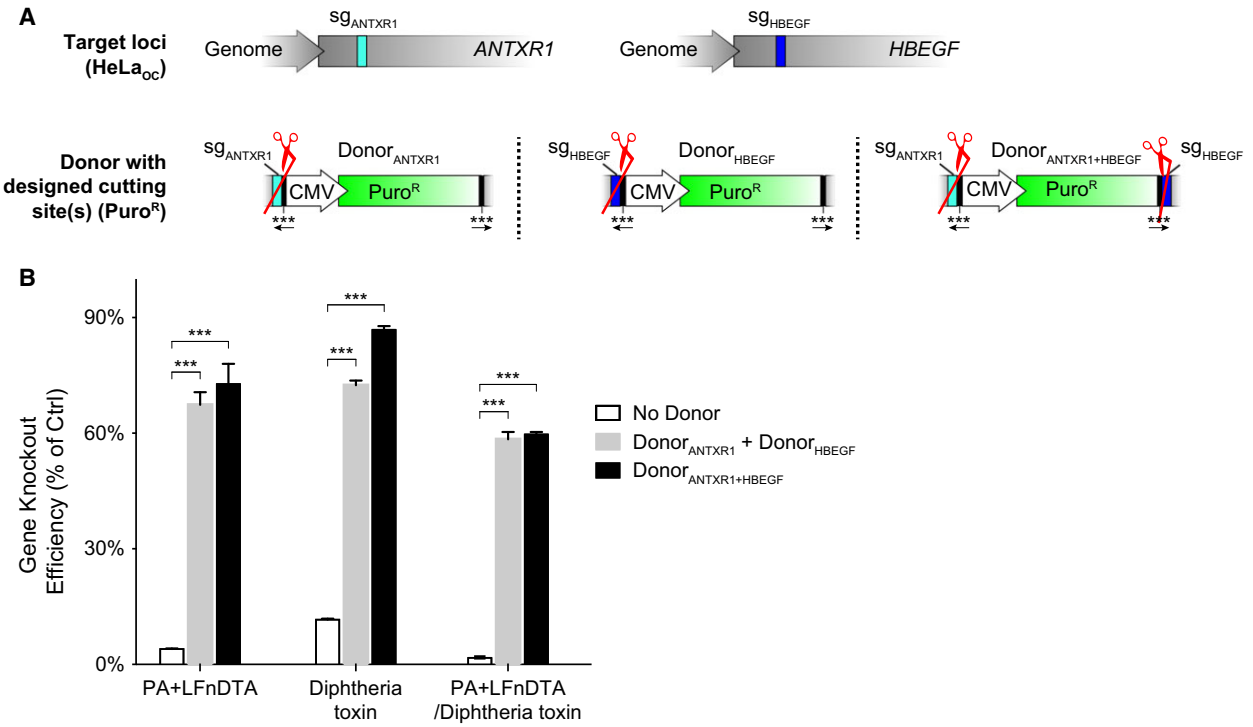


Fig. 4. One-step generation of *ANTXR1* and *HBEGF* knockouts in HeLa cells. (A) Schematic of the homologous recombination (HR)-independent knock-in of linear donor(s) on the pgRNA$_{ANTXR1+HBEGF}$-targeting sites of *ANTXR1* and *HBEGF* genes. Donor$_{ANTXR1}$ or Donor$_{HBEGF}$ has the sgRNA1$_{ANTXR1}$ or sgRNA2$_{HBEGF}$ cutting site at its 5′ end. Donor$_{ANTXR1+HBEGF}$ has the sgRNA1$_{ANTXR1}$ cutting site at the 5′ end and the sgRNA2$_{HBEGF}$ cutting site at the 3′ end. (B) *ANTXR1* and *HBEGF* knockout rate as a percentage of cells resistant to PA/LFnDTA or/and diphtheria toxin (DT). HeLa$_{OC}$ cells were transfected with pgRNA$_{ANTXR1+HBEGF}$ as indicated with corresponding donors or with a puromycin-resistant plasmid. Cells were selected with puromycin (1 μg·mL^{-1}) prior to the PA/LFnDTA and DT resistance analysis. Error bars indicate SD ($n = 3$), *t*-test, $^{**}P < 0.01$, $^{***}P < 0.001$.

designed, and their expression plasmid (pgRNA$_{PSEN}$, expressing both sgRNA$_{PSEN1}$ and sgRNA$_{PSEN2}$) and corresponding donors were created. We designed two kinds of donors as follows: one with two separate donors (Donor$_{PSEN1}$ and Donor$_{PSEN2}$), each harbouring one sgRNA recognition site, and the other donor (Donor$_{PSEN}$) with two-sgRNA recognition sites at each end (Fig. 3A). Notably, pgRNA$_{PSEN}$, or individual sgRNAs targeting either *PSEN1* or *PSEN2*, only exhibited mediocre activity in the T7E1 assay [26] in generating indels at two specific sites (Fig. S7A). Encouragingly, we were still able to obtain puro$^+$ clones using either Donor$_{PSEN}$ or Donor$_{PSEN1}$+ Donor$_{PSEN2}$ (Fig. S7B). Two clones identified from each case were selected for genome sequencing analysis (Fig. S7C). Both clones exhibited disruption of both *PSEN1* and *PSEN2* (Fig. 3B,C), demonstrating the superiority of this new method in generating gene knockouts.

To further compare the efficiency of our approach with the conventional method in generating multi-gene knockouts in a quantitative manner, we chose to simultaneously knockout both *ANTXR1* and *HBEGF* genes in HeLa cells. We designed two kinds of donors as follows: one with two separate donors (Donor$_{ANTXR1}$ + Donor$_{HBEGF}$), each harbouring an sgRNA site, and the other donor (Donor$_{ANTXR1+HBEGF}$) with two sgRNA cutting sites at each end (Fig. 4A). As described above, we were able to readily calculate the percentages of single knockouts (*ANTXR1* or *HBEGF*) and double knockouts (*ANTXR1* and *HBEGF*) from the portions of cells that were resistant to either PA/LFnDTA or DT, and to both toxins. Notably, we observed much improved efficiency using donor selection (either one of two donors) for double knockouts by a factor of approximately 30-fold in comparison with donor-free method (Fig. 4B). Consistent with our prior observation, the efficiency to obtain single knockouts was also greatly improved with donor enrichment (Fig. 4B).

We continued to test the limits of our method by targeting multiple genes at one time. We chose the *HSPA* gene family, whose members share sequence

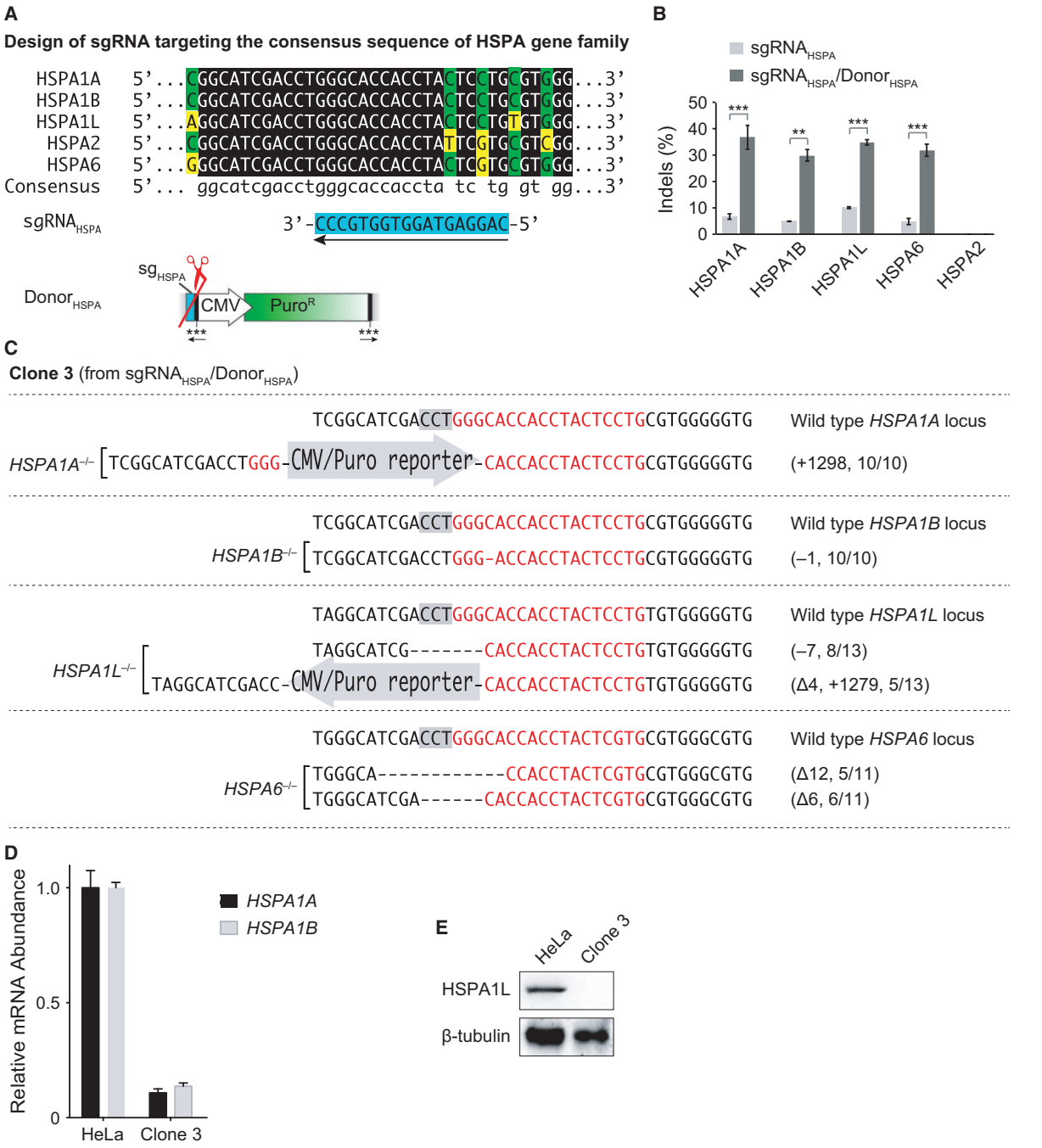


Fig. 5. One-step generation of multi-gene knockouts in HeLa cells. (A) Design of the sgRNA targeting the consensus sequence of five HSPA family genes and the universal linear donor ($Donor_{HSPA}$) used for the enrichment of cells containing multi-gene mutations. The consensus sequence of the HSPA gene family was analysed by multiple sequences alignment. The black-shaded nucleotides represent the consensus sequence of all five HSPA genes. The green-shaded nucleotides represent the consensus sequence of three or four *HSPA* genes, and the yellow-shaded nucleotides represent the nonconsensus nucleotides. (B) $sgRNA_{HSPA}$-triggered indels on five target genes in the absence and presence of $Donor_{HSPA}$, after puromycin selection. Indel efficiencies were evaluated with a T7E1 assay. Error bars indicate SD ($n = 3$), *t*-test, $^{**}P < 0.01$, $^{***}P < 0.001$. (C) Partial coding sequences of *HSPA1A*, *HSPA1B*, *HSPA1L* and *HSPA6* genes in the genome of HeLa clone 3 containing the sgRNA-targeting regions (red) and sequencing analysis of the mutated alleles. Clone 3 was from $HeLa_{OC}$ cells transfected with $sgRNA_{HSPA}/Donor_{HSPA}$. Shaded nucleotides represent the PAM sequence, and the dashes indicate deletions. Light-grey arrows in the background indicate the CMV promoter direction in the donor. (D) mRNA levels of *HSPA1A* and *HSPA1B* (normalized to *GAPDH*) of clone 3 was quantified by real-time PCR analysis. (E) Western blotting analysis for HSPA1L and β-tubulin (loading control) of clone 3.

homology. We purposefully selected one specific sgRNA that simultaneously targets three genes, *HSPA1A*, *HSPA1B*, *HSPA1L*, because of sequence identity. This sgRNA-targeting sequence has one mismatch at *HSPA6* loci and two mismatches at *HSPA2* loci. The linear donor was generated accordingly through PCR amplification (Fig. 5A). We applied the same protocol targeting the *HSPA* gene family in $HeLa_{OC}$. To compare the mutation rates of multi-gene targeting in the absence and presence of the donor reporter, we performed the T7E1 assay at all five gene loci, *HSPA1A*, *HSPA1B*, *HSPA1L*, *HSPA2* and *HSPA6*. In comparison to conventional gene knockout protocols (transfection with $sgRNA_{HSPA}$ alone), the usage of $Donor_{HSPA}$ increased the mutation rates by approximately 5.5-fold at the *HSPA1A* locus, 6.1-fold at the *HSPA1B* locus, 3.4-fold at the *HSPA1L* locus and 6.6-fold at the *HSPA6* locus (Fig. 5B). Interestingly, no detectable indels were observed at the *HSPA2* locus with or without the donor reporter, indicating that two mismatches completely abolished $sgRNA_{HSPA}$ recognition and, more importantly, the donor selection did not increase the risk of off-target effects (Fig. 5B). The cell pool sequencing results with or without donor transfection at *HSPA* family gene loci were consistent with the results of the T7E1 assay (Fig. S8). Notably, the T7E1 assay demonstrated that the probability to acquire the four-gene mutation has been increased by a factor of approximately 753 ($5.5 \times 6.1 \times 3.4 \times 6.6$)-fold in comparison with the conventional method without a donor. Considering that this calculation does not include those loci with donor insertions, the real efficiency increase was even higher. We then performed PCR verification at the five targeted loci of the puromycin-resistant single clones, and six clones identified using our new protocol that carry donor inserts in at least two target genes were selected for genome sequence analysis (Fig. S9). Clone 3 had modifications in four-gene loci: *HSPA1A*, *HSPA1B* and *HSPA1L,* with frame-shift mutations generating complete knockouts, and *HSPA6,* with two in-frame mutations (Fig. 5C). Sanger sequencing revealed that both *HSPA1A* and *HSPA1B* genes harbour premature termination (nonsense) codons because of donor insertion or sgRNA-mediated indels in clone 3. We predicted that mRNAs of these two genes might be subjected to nonsense-mediated mRNA decay (NMD) [36,37]. Indeed, expression of *HSPA1A* and *HSPA1B* of clone 3 was drastically reduced (Fig. 5D). In addition, HSPA1L production was undetectable in clone 3 from western blotting analysis (Fig. 5E). These results demonstrated that our method using specially designed donors enables the generation of multi-gene knockouts in human cancerous cells, a daunting task that is not possible with current methods.

In summary, conventional gene knockout strategies are usually labour-intensive and time-consuming because they lack effective stimulation and enrichment for clones containing target gene modifications [22]. The reported HR-independent linear donor insertion strategies [34,35,38,39] mostly rely on in-frame knock-in of exogenous gene, or restoration of reading frame of a reporter. Our redesigned protocol uses a linear donor with features specifically for gene knockouts, especially that the selection marker is driven by its own CMV promoter to ensure its high expression regardless of native promoter strength and open reading frame of target gene. The linear donor could be readily obtained through PCR amplification, a much easier step than plasmid donor construction. Importantly, we found that cleavage of donor fragments is essential for its effective integration because dummy donors without sgRNA recognition sites failed to promote donor integration (Fig. 1B). We reason that the abundant presence of DSBs generated by sgRNA-mediated cleavage of donors boosts the NHEJ activity, greatly stimulating DNA DSB rejoining at the target loci on chromosomes. In addition, it has been reported that multiple DSBs could accumulate to the pre-existing DNA damage repair centre and consequently accelerate the damage repair [40–42] and the incorporation of donors with DSBs. It becomes evident that as the sgRNA efficiency decreases or the number of target genes for disruption increases, higher fold of enrichment is observed with this strategy. These coordinate changes make the new approach extremely useful for targeting genes for which designing sgRNAs is difficult and in cases where several genes need to be targeted simultaneously. We anticipate that this method will facilitate broader application of the CRISPR system in the biomedical study of genes and their functions.

Acknowledgements

We acknowledge the staff of the BIOPIC sequencing facility (Peking University) for their assistance, and National Center for Protein Sciences Beijing (Peking University) for help in fluorescence-activated cell sorting. The project was supported by funds from the National Science Foundation of China (NSFC31430025, NSFC31170126, NSFC81471909), Beijing Advanced Innovation Center for Genomics at Peking University and the Peking-Tsinghua Center for Life Sciences (to W.W.).

Author contributions

WW conceived of and supervised the project. YZ and HZ performed the experiments. YZ, HZ and WW wrote the manuscript.

References

1 Kim YG, Cha J and Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A* **93**, 1156–1160.

2 Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A and Bonas U (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–1512.

3 Moscou MJ and Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501.

4 Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ *et al.* (2011) A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* **29**, 143–148.

5 Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA and Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821.

6 Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE and Church GM (2013) RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826.

7 Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA *et al.* (2013) Multiplex genome engineering using CRISPR/Cas Systems. *Science* **339**, 819–823.

8 Phillips ER and McKinnon PJ (2007) DNA double-strand break repair and development. *Oncogene* **26**, 7799–7808.

9 Chapman JR, Taylor MR and Boulton SJ (2012) Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell* **47**, 497–510.

10 Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA *et al.* (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451.

11 Ma H, Naseri A, Reyes-Gutierrez P, Wolfe SA, Zhang S and Pederson T (2015) Multicolor CRISPR labeling of chromosomal loci in human cells. *Proc Natl Acad Sci U S A* **112**, 3002–3007.

12 Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS *et al.* (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479–1491.

13 Li HL, Gee P, Ishida K and Hotta A (2015) Efficient genomic correction methods in human iPS cells using CRISPR-Cas9 system. *Methods* **101**, 27–35.

14 Savic N and Schwank G (2016) Advances in therapeutic CRISPR/Cas9 genome editing. *Transl Res* **168**, 15–21.

15 Miyaoka Y, Chan AH, Judge LM, Yoo J, Huang M, Nguyen TD, Lizarraga PP, So PL and Conklin BR (2014) Isolation of single-base genome-edited human iPS cells without antibiotic selection. *Nat Methods* **11**, 291–293.

16 Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK and Sander JD (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**, 822–826.

17 Zhou Y, Zhu S, Cai C, Yuan P, Li C, Huang Y and Wei W (2014) High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature* **509**, 487–491.

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

19 Liao S, Tammaro M and Yan H (2015) Enriching CRISPR-Cas9 targeted cells by co-targeting the HPRT gene. *Nucleic Acids Res* **43**, e134.

20 Kim H, Um E, Cho SR, Jung C, Kim H and Kim JS (2011) Surrogate reporters for enrichment of cells with nuclease-induced mutations. *Nat Methods* **8**, 941–943.

21 Ramakrishna S, Cho SW, Kim S, Song M, Gopalappa R, Kim JS and Kim H (2014) Surrogate reporter-based enrichment of cells containing RNA-guided Cas9 nuclease-induced mutations. *Nat Commun* **5**, 3378.

22 Wang T, Wei JJ, Sabatini DM and Lander ES (2014) Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84.

23 Yuan P, Zhang H, Cai C, Zhu S, Zhou Y, Yang X, He R, Li C, Guo S, Li S *et al.* (2015) Chondroitin sulfate proteoglycan 4 functions as the cellular receptor for *Clostridium difficile* toxin B. *Cell Res* **25**, 157–168.

24 Yang J, Yuan P, Wen D, Sheng Y, Zhu S, Yu Y, Gao X and Wei W (2013) ULtiMATE system for rapid assembly of customized TAL effectors. *PLoS One* **8**, e75649.

25 Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, Wang N, Lee G, Bartsevich VV, Lee YL *et al.* (2008) Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol* **26**, 808–816.

26 Kim HJ, Lee HJ, Kim H, Cho SW and Kim JS (2009) Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res* **19**, 1279–1288.

27 Lackner DH, Carre A, Guzzardo PM, Banning C, Mangena R, Henley T, Oberndorfer S, Gapp BV,

Nijman SM, Brummelkamp TR *et al.* (2015) A generic strategy for CRISPR-Cas9-mediated gene tagging. *Nat Commun* **6**, 10237.
28 Auer TO and Del Bene F (2014) CRISPR/Cas9 and TALEN-mediated knock-in approaches in zebrafish. *Methods* **69**, 142–150.
29 Li K, Wang G, Andersen T, Zhou P and Pu WT (2014) Optimization of genome engineering approaches with the CRISPR/Cas9 System. *PLoS One* **9**, e105779.
30 Peng J, Zhou Y, Zhu S and Wei W (2015) High-throughput screens in mammalian cells using the CRISPR-Cas9 system. *FEBS J* **282**, 2089–2096.
31 Potter CJ and Luo L (2010) Splinkerette PCR for mapping transposable elements in *Drosophila*. *PLoS One* **5**, e10168.
32 Uren AG, Mikkers H, Kool J, van der Weyden L, Lund AH, Wilson CH, Rance R, Jonkers J, van Lohuizen M, Berns A *et al.* (2009) A high-throughput splinkerette-PCR method for the isolation and sequencing of retroviral insertion sites. *Nat Protoc* **4**, 789–798.
33 Yin B and Largaespada DA (2007) PCR-based procedures to isolate insertion sites of DNA elements. *Biotechniques* **43**, 79–84.
34 Nakade S, Tsubota T, Sakane Y, Kume S, Sakamoto N, Obara M, Daimon T, Sezutsu H, Yamamoto T, Sakuma T *et al.* (2014) Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. *Nat Commun* **5**, 5560.
35 Sakuma T, Nakade S, Sakane Y, Suzuki KT and Yamamoto T (2016) MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. *Nat Protoc* **11**, 118–133.
36 Brogna S and Wen J (2009) Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat Struct Mol Biol* **16**, 107–113.
37 Chang YF, Imam JS and Wilkinson MF (2007) The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem* **76**, 51–74.
38 Cristea S, Freyvert Y, Santiago Y, Holmes MC, Urnov FD, Gregory PD and Cost GJ (2013) In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration. *Biotechnol Bioeng* **110**, 871–880.
39 Orlando SJ, Santiago Y, DeKelver RC, Freyvert Y, Boydston EA, Moehle EA, Choi VM, Gopalan SM, Lou JF, Li J *et al.* (2010) Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Res* **38**, e152.
40 Costes SV, Ponomarev A, Chen JL, Nguyen D, Cucinotta FA and Barcellos-Hoff MH (2007) Image-based modeling reveals dynamic redistribution of DNA damage into nuclear sub-domains. *PLoS Comput Biol* **3**, e155.
41 DiBiase SJ, Zeng ZC, Chen R, Hyslop T, Curran WJ Jr and Iliakis G (2000) DNA-dependent protein kinase stimulates an independently active, nonhomologous, end-joining apparatus. *Cancer Res* **60**, 1245–1253.
42 Lisby M, Mortensen UH and Rothstein R (2003) Colocalization of multiple DNA double-strand breaks at a single Rad52 repair centre. *Nat Cell Biol* **5**, 572–577.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Appendix S1. Supplementary sequences.
Fig. S1. Enrichment of *ANTXR1* knockout cells in a pooled population and single clones through donor-mediated puromycin resistance selection.
Fig. S2. Puromycin selection-mediated enrichment of *HBEGF* disruption in HeLa cells.
Fig. S3. EGFP selection-mediated enrichment of *HBEGF* disruption in HEK293T cells.
Fig. S4. Puromycin selection-mediated enrichment of *ANTXR1* disruption in $HeLa_{OC}$ cells.
Fig. S5. Puromycin selection-mediated enrichment of *HBEGF* disruption in $HeLa_{OC}$ cells.
Fig. S6. Junction sequences of donor integration sites, *ANTXR1* locus 1 (A), *ANTXR1* locus 2 (B), *HBEGF* locus 1 (C) and *HBEGF* locus 2 (D) in $HeLa_{OC}$ cells.
Fig. S7. *PSEN1* and *PSEN2* sgRNA efficiency evaluation and single clone identification in $HeLa_{OC}$ cells.
Fig. S8. Sequencing chromatogram of targeted regions of the *HSPA* family genes from pooled cells with or without donor transfection.
Fig. S9. Single clone identification for the donor insertion at the targeting loci of five HSPA family genes, *HSPA1A*, *HSPA1B*, *HSPA1L*, *HSPA6* and *HSPA2*.
Table S1. T7E1 assay results of sgRNAs targeting *ANTXR1*, *HBEGF*, *PSEN1* and *PSEN2* in $HeLa_{OC}$ cells.
Table S2. Primers for sgRNA or pgRNA construction.
Table S3. Primers used for PCR amplification of the targeted genome sequences.
Table S4. PCR primers for linear donor construction.
Table S5. Primers used for splinkerette PCR.
Table S6. Multi-gene knockout summary.