Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology) 2022 23(2):141-152
www.jzus.zju.edu.cn; www.springer.com/journal/11585
E-mail: jzus_b@zju.edu.cn

Research Article
https://doi.org/10.1631/jzus.B2100196

A high-efficiency and versatile CRISPR/Cas9-mediated HDR-based biallelic editing system

Xinyi LI[1,2], Bing SUN[1], Hongrun QIAN[1], Jinrong MA[1], Magdalena PAOLINO[2], Zhiying ZHANG[1]✉

[1]*College of Animal Science and Technology, Northwest A&F University, Yangling 712100, China*
[2]*Karolinska Institute, Department of Medicine Solna, Center for Molecular Medicine, Karolinska University Hospital Solna, 17176 Stockholm, Sweden*

Abstract: Clustered regulatory interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 nuclease (Cas9), the third-generation genome editing tool, has been favored because of its high efficiency and clear system composition. In this technology, the introduced double-strand breaks (DSBs) are mainly repaired by non-homologous end joining (NHEJ) or homology-directed repair (HDR) pathways. The high-fidelity HDR pathway is used for genome modification, which can introduce artificially controllable insertions, deletions, or substitutions carried by the donor templates. Although high-level knock-out can be easily achieved by NHEJ, accurate HDR-mediated knock-in remains a technical challenge. In most circumstances, although both alleles are broken by endonucleases, only one can be repaired by HDR, and the other one is usually recombined by NHEJ. For gene function studies or disease model establishment, biallelic editing to generate homozygous cell lines and homozygotes is needed to ensure consistent phenotypes. Thus, there is an urgent need for an efficient biallelic editing system. Here, we developed three pairs of integrated selection systems, where each of the two selection cassettes contained one drug-screening gene and one fluorescent marker. Flanked by homologous arms containing the mutated sequences, the selection cassettes were integrated into the target site, mediated by CRISPR/Cas9-induced HDR. Positively targeted cell clones were massively enriched by fluorescent microscopy after screening for drug resistance. We tested this novel method on the amyloid precursor protein (*APP*) and presenilin 1 (*PSEN1*) loci and demonstrated up to 82.0% biallelic editing efficiency after optimization. Our results indicate that this strategy can provide a new efficient approach for biallelic editing and lay a foundation for establishment of an easier and more efficient disease model.

Key words: Biallelic editing; CRISPR/Cas9; Homology-directed repair (HDR); Homozygote

1 Introduction

The discovery and application of artificial endonucleases, including zinc finger nucleases (Bibikova et al., 2003), transcription activator-like effector nucleases (Boch et al., 2009), and clustered regulatory interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems (Jinek et al., 2012; Cong et al., 2013), have made arbitrary genomic modification a reality. To accurately modify the target sequence by genome editing, CRISPR-mediated base editing and prime-editing technologies were recently generated (Li et al., 2021); they can achieve targeted DNA insertions, deletions, and base conversions without generating double-strand breaks (DSBs) or needing donor DNA templates (as in most traditional genome-editing methods). However, for accuracy and efficiency, extensive knowledge of the precise design is needed to avoid off-target base editing and produce appropriate prime-editing guide RNA (pegRNA) designs (Anzalone et al., 2019; Zuo et al., 2019). Currently, homology-directed repair (HDR)-based CRISPR/Cas9 methods are still the most used due to their high efficiency and simple design principle (Mali et al., 2013; Hsu et al., 2014). Plasmids (Koch et al., 2018), polymerase chain reaction (PCR) fragments (Chu et al., 2015; Shao et al., 2017), and single-stranded oligonucleotides (ssODNs) (Yang et al., 2013; Paquet et al., 2016) are frequently used as donors, to introduce base insertions, deletions, or substitutions. Without selection markers,

✉ Zhiying ZHANG, zhangzhy@nwafu.edu.cn
Zhiying ZHANG, https://orcid.org/0000-0001-6012-0889

Received Feb. 27, 2021; Revision accepted June 20, 2021;
Crosschecked Dec. 15, 2021

HDR efficiency only reaches a 15% positive rate at the cell line level and 18% for zygotes (Paquet et al., 2016; Codner et al., 2018). Its efficiency further decreases as the distance between the target site and introduced mutation increases (Paquet et al., 2016). Moreover, detection of positive clones is in general time-consuming and labor-intensive. To efficiently improve the acquisition of positive cell clones, plasmid donors containing genes resistant to drug screening can be used. However, the efficiency of correct and simultaneous modification of both alleles remains a challenge. Biallelic editing of homozygous cell lines and model animals is a reliable tool to study gene functions and species characteristics, as they have stable phenotypes and do not produce trait segregation. Thus, efficient biallelic gene editing is essential for gene function research.

The introduction of drug-screening genes or fluorescent protein genes into reporter vectors or genomes can significantly improve the positive rate of targeted cells (Ren et al., 2015; Wang et al., 2017; Wu et al., 2017; Yan et al., 2020). However, single selection systems cannot discriminate between monoallelic and biallelic targeting. Therefore, we hypothesized that using two donor vectors, each containing a different fluorescent protein as well as a drug-resistant gene, could lead to higher biallelic editing efficiency. Here, we have developed three pairs of such donors for targeted allele integration mediated by CRISPR/Cas9. All three were proved to be efficient and versatile, allowing a highly efficient selection of biallelic targeted clones.

2 Materials and methods

2.1 Preparation for genome targeting

The targeting plasmid was constructed by inserting the single-guide RNA (sgRNA) fragment under the human U6 promoter on the pLL3.7 backbone, which also contained the *Streptococcus pyogenes* Cas9 protein expression sequence. The efficiency of the targeting plasmids was detected using the single-strand annealing (SSA)-DsRed-puromycin resistance gene (Puro)-enhanced green fluorescent protein (eGFP) (RPG) surrogate reporter, with the insertion of the target sequence amplified by target-F/target-R (Ren et al., 2015). The primers used for genome targeting preparation can be found in Table 1.

To ensure that both alleles could be targeted by CRISPR/Cas9, sequences of the target sites were amplified by a target-F/target-R primer pair and confirmed in advance by Sanger sequencing. The primers used in homozygosity detection are listed in Table 1.

2.2 Construction of the selection cassette and donor vectors of the TK-Puro-eGFP/TK-Zeo-mRFP (TPG/TZR) system

Firstly, the eGFP-bovine growth hormone (BGH) polyA module in the JMB81-thymidine kinase (TK)-Puro-T2A-eGFP backbone (used for the SSA-mediated precise gene-editing system (Li et al., 2018)) was replaced by a monomeric red fluorescent protein (mRFP)-BGH polyA fragment and it was cloned from plasmid pB-Puro-mRFP following *Nhe*I and *Xho*I enzymatic digestion. Next, the phosphoglycerate kinase (PGK) promoter-zeocin resistance gene (Zeo)-T2A module was cloned by overlap PCR. The PGK promoter was amplified from the JMB81-TK-Puro-T2A-eGFP backbone using PGK-F/R primer pair. The first half of the T2A sequence was attached to the end of Zeo by amplification using Zeo-F and Zeo-T2A-R1 primers. The PGK promoter and Zeo-T2A were then linked using PGK-F and Zeo-T2A-R2 primers, with the rest half of the T2A sequence added simultaneously. We chose the amyloid precursor protein (*APP*) gene locus to detect the feasibility of our TPG/TZR biallelic editing system because the *APP*.Donor (TPG) vector containing the *APP* homologous arms was available in our lab from previous research (Li et al., 2018). To generate the double-cut donor, the *APP* target sequence (GGGTTGACAAATATCAAGACGG) was added to the 5' and 3' of the homologous arms by PCR, using the following primers: APP-Left arm-F (*Eco*RI) and APP-Right arm-R (*Cla*I). Finally, the TPG selection cassette in the *APP*. Donor (TPG) (Fig. S1a) was replaced by the TZR selection cassette cut from the JMB81-TK-Zeo-T2A-mRFP backbone to construct the *APP*.Donor (TZR) (Fig. S1b).

The homologous-arm plasmids for targeting the presenilin 1 (*PSEN1*) locus were constructed according to the same strategy as the *APP* locus (Li et al., 2018). The *PSEN1* target sequence (TGTTGTCATGACTATC CTCCTGG) was added to the 5' and 3' of the homologous arms by PCR primers: PSEN1. LA-F (*Bam*HI) and PSEN1.RA-R (*Hin*dIII). Then, the TPG or TZR cassette was inserted into the homologous plasmid to

Table 1 Primers used for constructing donor vectors and for PCR detection

Primer name	Sequence (5'→3')
U6-F (*Xba*I)	cgc*TCTAGA*TTTCCCATGATTCCTTCATAT
U6-R (*Not*I)	acc*GCGGCCGC*AAAAAAGCACCGACTCGGTGCC
APP-Left arm-F (*Eco*RI)	ccg*GAATTC***GGGTTGACAAATATCAAGACGG**CAAAGGATGGAAGTTACAGG
APP-Left arm-R (*Not*I)	aat*GCGGCCGC*CGCTAAAGAGGAAGAGGACAGGGTCAGTCTTGATATTTGTCAACCCAG
APP-SSA-F (*Xho*I)	ccg*CTCGAG*CGCTAAAGAGGAAGAGGACAGGGAAAGCCCCAAAGTAGCAGTT
APP-Right arm-R (*Cla*I)	ccc*ATCGAT***GGGTTGACAAATATCAAGACGG**TTAACCCTCTTACTGCACCT
PSEN1.target-F	aat*GCGGCCGC*TCACAGAAGATACCGAGACT
PSEN1.target-R	cgc*GGATCC*GACAAGAATACCCAACCATAAG
PSEN1.sgRNA-F	TGTTGTCATGACTATCCTCCGTTTTAGAGCTAGAAATAGCAAG
PSEN1.sgRNA-R	GGAGGATAGTCATGACAACAGGTGTTTCGTCCTTTCCACAAG
PSEN1.LA-F (*Bam*HI)	cgc*GGATCC***TGTTGTCATGACTATCCTCCTGG**TCTGACAGTAGGAAGGAGTAGG
PSEN1.LA-R (*Not*I)	aat*GCGGCCGC*CGCTAAAGAGGAAGAGGACAGGGCACGACAACAATGACACTGATCATG
PSEN1.SSA-F (*Xho*I)	ccg*CTCGAG*CGCTAAAGAGGAAGAGGACAGGGACCTTTCTTCTGACAACCACTT
PSEN1.SSA-R	TCACGACAACAATGACACTGATCATG
PSEN1.RA-F	ATCAGTGTCATTGTTGTCGTGAGGATCCTCCTTGTGGTTCTGTATAAATACAG
PSEN1.RA-R (*Hin*dIII)	ccc*AAGCTT***TGTTGTCATGACTATCCTCCTGG**TAGGCAAGGTGTGACCAGTC
pGK-F	cgc*GGATCC*CCGGTAGGCGCCAA
pGK-R	TCAACTTGGCCATATTGGCTGCAGGTCGAA
Zeo-F	AGCCAATATGGCCAAGTTGACCAGTGCCGT
Zeo-T2A-R1	ACCGCATGTTAGCAGACTTCCTCTGCCCTCGTCCTGCTCCTCGGCCACGA
Zeo-T2A-R2	cac*GCTAGC*TGGGCCAGGATTCTCCTCGACGTCACCGCATGTTAGCAGACTTC
P1	GGCTTGCTCTGTGTTGAATC
P2	TCCCTATTGGCGTTACTATGG
P3	TGAGCAAAGACCCCAACGAG
P3-2	CTACGACGCCGAGGTCAAGA
P4	GCAAATCCAAATCCCAAGTCAG
P5	CCAGATTAGTACGGTGGCTTCA
P6	ACAGTGGAGATGGTGAGAGGAT
P2A-F1 (*Eco*RI)	ccg*GAATTC*GCAACAAACTTCTCACTA
P2A-R	ACGAGGCCATAGGCCCGGGATTCTCCTCCA
TK-F	TCCCGGGCCTATGGCCTCGTACCCCGGCCA
TK-R (*Xho*I)	CCGCTCGAGCAGACATGATAAGATACA
P2A-F2 (RFP)	CACCGGCACCGCAACAAACTTCTCACTACT
RFP-F (*Nhe*I)	CTAGCTAGCATGGCCTCCTCCGAGGAC
RFP-R	AGTTTGTTGCGGTGCCGGTGGAGTGGCGGC
TK-F2 (*Nhe*I)	CTAGCTAGCATGGCCTCGTACCCCGGCCA
TK-R2 (*Eco*RI, *Pme*I)	CCG*GAATTCGTTTAAAC*CTA*ACCGGT*GTTAGCCTCCC
P7	GCTTGTACTCGGTCATATTGG
P7-2	AACGGCACTGGTCAACTTGG
P8	CCAACGGCGACCTGTATAAC

Lower case: protective base for enzyme site; Italics: enzyme site; Bold: CRISPR/Cas9 target site. PCR: polymerase chain reaction; CRISPR: clustered regulatory interspaced short palindromic repeats; Cas9: CRISPR-associated protein 9 nuclease.

produce *PSEN1*. Donor (TPG) (Fig. S1c) or *PSEN1*. Donor (TZR) vectors (Fig. S1d). All primers used to construct the donor plasmids are listed in Table 1.

2.3 Cell culture and transfection

HEK293T cells were used for all transfection in this study and were maintained in Dulbecco's modified Eagle's medium (DMEM; Gibco, USA) supplemented with 100 U/mL penicillin, 100 μg/mL streptomycin, and 10% (volume fraction) fetal bovine serum (FBS; HyClone, USA), at 37 ℃ and 5% CO_2. HEK293T cells were seeded into 24-well plates at a density of 40% and were transfected using Lipofectamine™ 2000 (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's manual, when cell confluence reached 90%.

2.4 Drug screening, expansion, and detection of cell clones

One day (24 h) after transfection, one quarter of the cells were transferred into a 100-mm plate and allowed to grow for an additional 5 d. Six days after

transfection, the medium was supplemented with puromycin (3 μg/mL; Invitrogen) and zeocin (600 μg/mL; Invitrogen). After 3 d of drug treatment, the concentration of puromycin (Invitrogen) was reduced to 1 μg/mL; the concentration of zeocin remained unchanged. After another 7 d of selection, and upon control confirmation that all negative clones had died, we removed drugs from the media to minimize their negative effects on the growth of positive colonies and prevent random insertion caused by drug selection pressure. A number of cell colonies (about 50) with fluorescence were picked and seeded into 24-well plates for proliferation. Three-quarters of the cells for each clone were used for genomic DNA preparation using the E.Z.N.A. DNA Kit from OMEGA Bio-Tek (GA, USA) along with PCR analysis to confirm the correct integration. The other quarter were left in the well to grow until confluency and then frozen.

To verify the insertion of the selection cassette, we designed primers located beside the homologous arm sequences. For the application of the 5' homologous arm, the forward primer was located upstream of the 5' homologous arm, and the reverse primer was located just upstream of the selection cassette (TPG/TZR, Puro-TK/Zeo-TK (PT/ZT), Puro-eGFP-TK/Zeo-mRFP-TK (PGT/ZRT)). For the application of the 3' homologous arm, the forward primer was located downstream of the selection cassette, whereas the reverse primer was located downstream of the 3' homologous arm. Primers flanking the homologous arms were used to detect the biallelic targeting clones. After integration of the selection cassette, the distance between the two flanking primers became too long to be amplified, so only the wild-type genome could produce the amplicons. All primer sequences are shown in Table 1.

2.5 Construction of PGT/ZRT selection cassette and donor vectors

The *TK* gene was cut off from the TPG/TZR selection-cassette plasmids (JMB81-TK-Puro-T2A-eGFP and JMB81-TK-Zeo-T2A-mRFP) by *Cla*I/*Bam*HI restriction enzyme digestion. Later, the JMB81-pPGK-Puro-T2A-eGFP-BGH polyA and JMB81-pPGK-Zeo-T2A-mRFP-BGH polyA vectors were produced after sticky end fill-in with *EasyPfu* DNA polymerase (Trangen Biotech, China) and T4 ligase ligation (NEB, USA).

The P2A shear peptide was cloned from the pU6-CBh-Cas9-T2A-mcherry-P2A-Ad4E4orf6 plasmid (Addgene #64222) using the primer pair P2A-F1 (*Eco*RI) and P2A-R. The TK-SV40T polyA fragment was obtained by amplification with the primer pair TK-F and TK-R (*Xho*I). P2A and TK-SV40T polyA were linked by overlap PCR using primer pair P2A-F1 (*Eco*RI) and TK-R2 (*Xho*I). Then the P2A-TK-SV40T polyA fragment was inserted into the intermediate JMB81-pPGK-Puro-T2A-eGFP-BGH polyA plasmid to generate the PGT selection cassette plasmid JMB81-pPGK-Puro-T2A-eGFP-P2A-TK-SV40T polyA.

The mRFP fragment was amplified by primer pair RFP-F (*Nhe*I) and RFP-R and later added to the 5' of the P2A-TK-SV40T polyA fragment, which was itself amplified by the primer pair P2A-F2 (RFP) and TK-R (*Xho*I). Later, the mRFP-P2A-TK-SV40T polyA fragment was inserted into the intermediate JMB81-pPGK-Zeo-T2A-mRFP-BGH polyA plasmid to produce the ZRT selection-cassette plasmid JMB81-pPGK-Zeo-T2A-mRFP-P2A-TK-SV40T polyA.

Lastly, the generated PGT/ZRT selection cassettes were transferred to the *APP*.Donor vectors by *Not*I and *Xho*I enzymatic restrictions, producing the final *APP*. Donor (PGT) and *APP*. Donor (ZRT) (Figs. S1e and S1f).

2.6 Construction of PT/ZT selection cassette and donor vectors

The *TK* gene fragment amplified by the primer pair TK-F2 (*Nhe*I) and TK-R2 (*Eco*RI, *Pme*I) and the *eGFP* gene of the JMB81-TK-Puro-T2A-eGFP plasmid replaced the *TK* gene by *Nhe*I/*Eco*RI digestion and ligase reaction. Simultaneously, the *mRFP* gene on the JMB81-TK-Zeo-T2A-mRFP plasmid was replaced by the *TK* gene using *Nhe*I/*Pme*I digestion. The newly constructed PT/ZT selection-cassette plasmids were then inserted into the *APP*.Donor (PT) (Fig. S1g) and *APP*. Donor (ZT) (Fig. S1h) by enzyme restriction with *Not*I and *Xho*I.

3 Results

3.1 Design and construction of the TPG/TZR for HDR-based biallelic editing of the *APP* locus

We chose the TK-Puro-eGFP backbone as one of the selection cassette systems (Fig. 1a), because we

had used it effectively in previous studies (Li et al., 2018). Zeocin and mRFP, which were learned from previous plasmid structures, were introduced (for fluorescence microscopy and drug screening, respectively) into a second selection cassette, TK-Zeo-mRFP, constructed from the TK-Puro-eGFP backbone (Fig. 1a). For both *APP.*Donors, a double-cut donor design was applied to reach higher HDR efficiency (Chen et al., 2017; Zhang et al., 2017), and synonymous mutations were introduced to create *Xba*I restriction sites for later detection and prevention of secondary targeting. The *TK* gene was included to allow removal of the selection cassette from the genome after specific modification. To achieve biallelic editing, we transfected the biallelic editing system, consisting of the three vectors *APP.*CRISPR/Cas9, *APP.*Donor (TPG), and *APP.*Donor (TZR) (Figs. S1a and S1b), into HEK293T cells (Fig. 1b). After 10-d dual drug screening for puromycin and zeocin resistance, 51 cell clones with both green and red fluorescence were successively picked, expanded, and analyzed by PCR and Sanger sequencing (Figs. 1c–1e). We used PCR with specific primer pairs for the 5' homologous arm (1315 bp, P1/P2), TPG 3' homologous arm (2225 bp, P3/P4), and TZR 3' homologous arm (2360 bp, P3-2/P4) to select the clones in which biallelic editing was successful (Figs. 1b and 1d). The PCR amplicon for the 3' homologous arm could be digested by the *Xba*I restriction endonuclease, confirming successful introduction of the point mutation in the *APP* locus (Fig. 1d). The insertion of the selection cassettes into both alleles was further confirmed by the fact that there was no band amplification as a result of primer pair P1/P4 flanking the homologous arm region; as the amplicon was too long to be amplified after the insertion of the selection cassette, only the wild-type allele can produce amplified products (Figs. 1b and 1d). Finally, all three homologous arm detection fragments were detected by Sanger sequencing, confirming effective editing on both alleles (Fig. 1e). Out of the 51 selected clones, 27 revealed biallelic editing at the *APP* locus; this represented a 52.94% positive biallelic targeting rate.

3.2 HDR-based biallelic editing of the *PSEN1* locus

To test the universality of our biallelic editing system, we adapted the TPG/TZR selection cassettes to edit the *PSEN1* locus, following the same design strategy as for the *APP* locus. Instead of *Xba*I, we introduced a *Bam*HI restriction site in the 3' homologous arm for detecting the point mutation after HDR (Fig. 2a). After the transfection of the *PSEN1*.CRISPR/Cas9, *PSEN1*.Donor (TPG), and *PSEN1*.Donor (TZR) plasmid mixtures (Figs. S1c and S1d), HEK293T cell clones with dual drug resistance and dual fluorescence were expanded for analysis. PCR amplification was used to detect the three homologous arms: 5' (1138 bp, P5/P2), TPG 3' (1947 bp, P3/P6), and TZR 3' (2082 bp, P3-2/P6). Primers flanking the TPG 3' arm and TZR 3' arm were digested by the *Bam*HI restriction endonuclease, verifying the success of biallelic editing (Fig. 2b). No evident amplification with the flanking primer pair (P5/P6), as well as the Sanger sequencing results, conclusively proved the biallelic integration of the selection cassette (Fig. 2c). In summary, 28 out of 51 picked clones had the correct biallelic editing on the *PSEN1* locus, representing a 54.90% positive rate.

3.3 Optimization of the biallelic editing system

Since our 5 kb-length TPG/TZR system gave us only approximately 50% positive rates at two different loci, we next tested to find out whether a shorter cassette would improve the efficiency of our biallelic editing system. For comparison, we designed two sets of shorter selection cassettes to target the *APP* locus again (Fig. 3a). We shortened the vectors to 3.2 and 2.9 kb (for PGT and ZRT, respectively) by using self-cleaving peptides for the expression of the *Puro*/*Zeo* and *eGFP*/*mRFP* screening genes simultaneously, with the same promoter and terminal sequences. Subsequently, to test the efficiency of the optimized biallelic editing systems, we cloned the newly constructed pairs of selection cassettes into the *APP.*Donor vectors (Figs. S1e and S1f) and co-transfected them with the targeting plasmid *APP.*CRISPR/Cas9 (Fig. 3b).

After drug screening, clones from the PGT/ZRT-targeted cells were expanded and detected by PCR amplification (Figs. 3c and 3d). To detect the two 5' homologous arms after biallelic editing, we combined reverse primers P7 and P7-2 (located in the 5' end of the *Puro*/*Zeo* sequences, respectively) with primer P1. The 3' homologous arm was amplified with the forward primer P8 (located in the 3' end of the *TK* gene) and primer P4, and by subsequent digestion of the amplicon with the *Xba*I endonuclease (Figs. 3b and 3d). Sanger sequencing results for the two 5' homologous

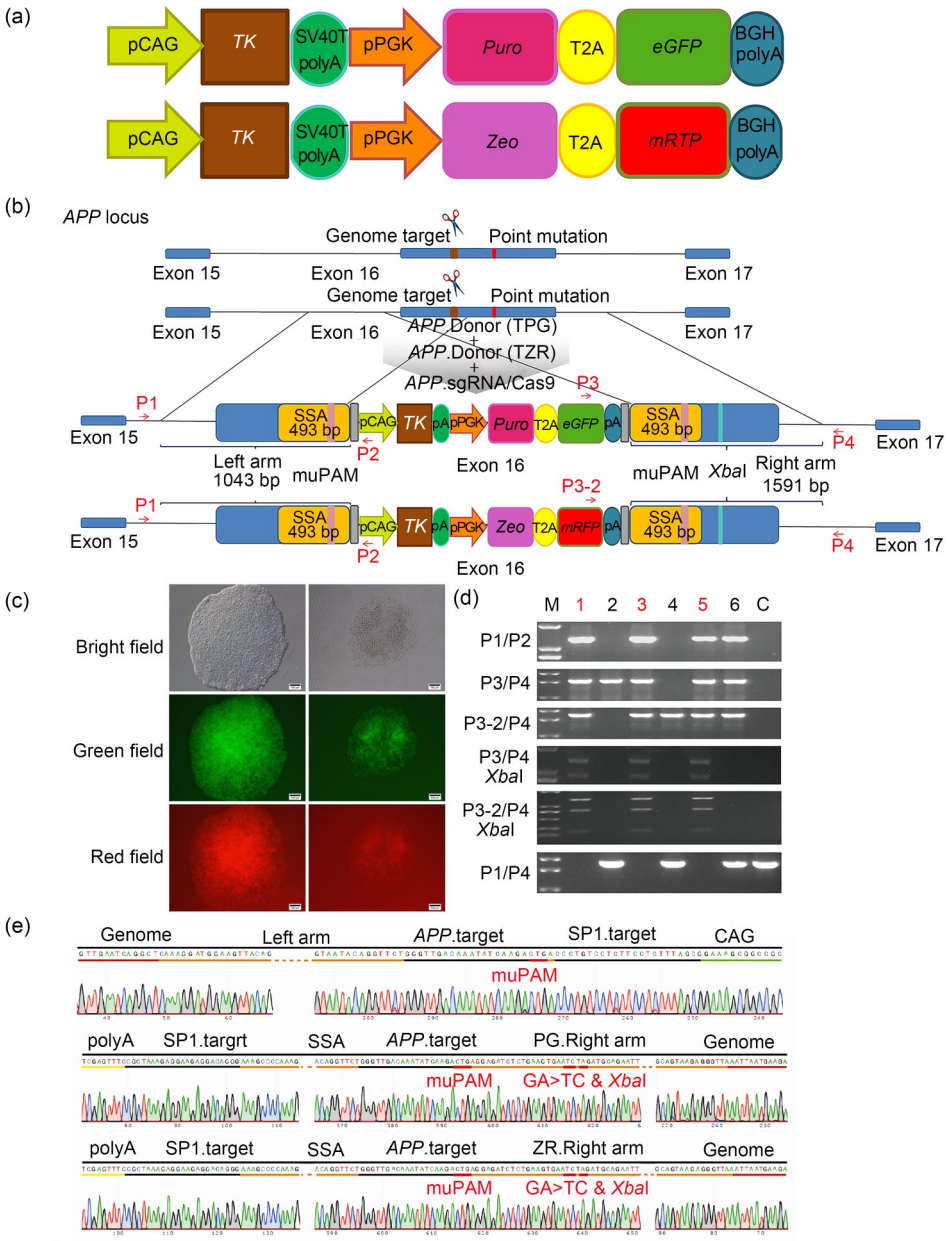


Fig. 1 TK-Puro-eGFP/TK-Zeo-mRFP (TPG/TZR)-dependent biallelic genome editing at the *APP* locus. (a) TPG/TZR selection cassettes. Two separate cassettes, each containing a drug-resistant and fluorescence marker, were used for allele integration, and were constructed using the same structure of promoters, self-cleaving peptides, and terminators. (b) Schematic diagram of exon 16 in the *APP* locus before (upper part) and after (lower part) insertion of the TPG/TZR cassettes and introduction of the point mutation according to homology-directed repair mediated by the CRISPR/Cas9 system. Red arrows indicate primers used for PCR analysis. (c) Representative fluorescent microscope photos of the positively selected HEK293T clones after 10-d puromycin and zeocin screening. At this point, clones were picked and seeded into 24-well plates for expansion and analysis. Scale bar=100 μm. (d) 5' and 3' recombinations and biallelic integrations were detected by PCR with primer pairs: P1/P2, P3/P4 (TPG), P3-2/P4 (TZR), and P1/P4, respectively. Point mutations on the P3/P4 (TPG) and P3-2/P4 (TZR) PCR fragments can be verified by *Xba*I restriction endonuclease reaction. C: control, a non-targeted HEK293T control sample; M: trans 2K plus DNA marker; Lanes 1 to 6: single colonies (red ones are positive colonies). (e) Sanger sequencing results of the PCR fragments amplified from the three homologous arms, confirming the desired point mutation at the *APP* locus. *APP*: amyloid precursor protein; PCR: polymerase chain reaction; CRISPR: clustered regulatory interspaced short palindromic repeats; Cas9: CRISPR-associated protein 9 nuclease; sgRNA: single-guide RNA; pCAG: cytomegalovirus (CMV) immediate enhancer/β-actin (CAG) promoter; *TK*: thymidine kinase; pA: polyadenylic acid; pPGK: phosphoglycerate kinase promoter; *Puro*: puromycin resistance gene; *Zeo*: zeocin resistance gene; *eGFP*: enhanced green fluorescent protein; *mRFP*: monomeric red fluorescent protein; BGH: bovine growth hormone; SSA: single-strand annealing; muPAM: mutated PAM.

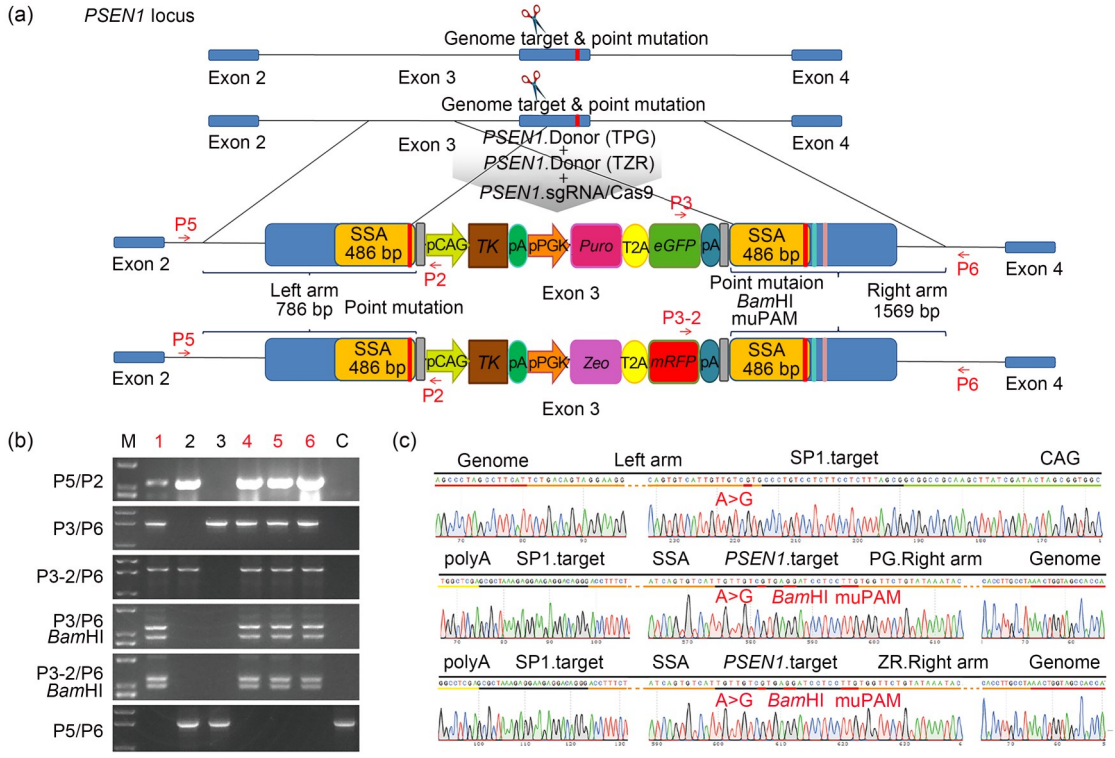


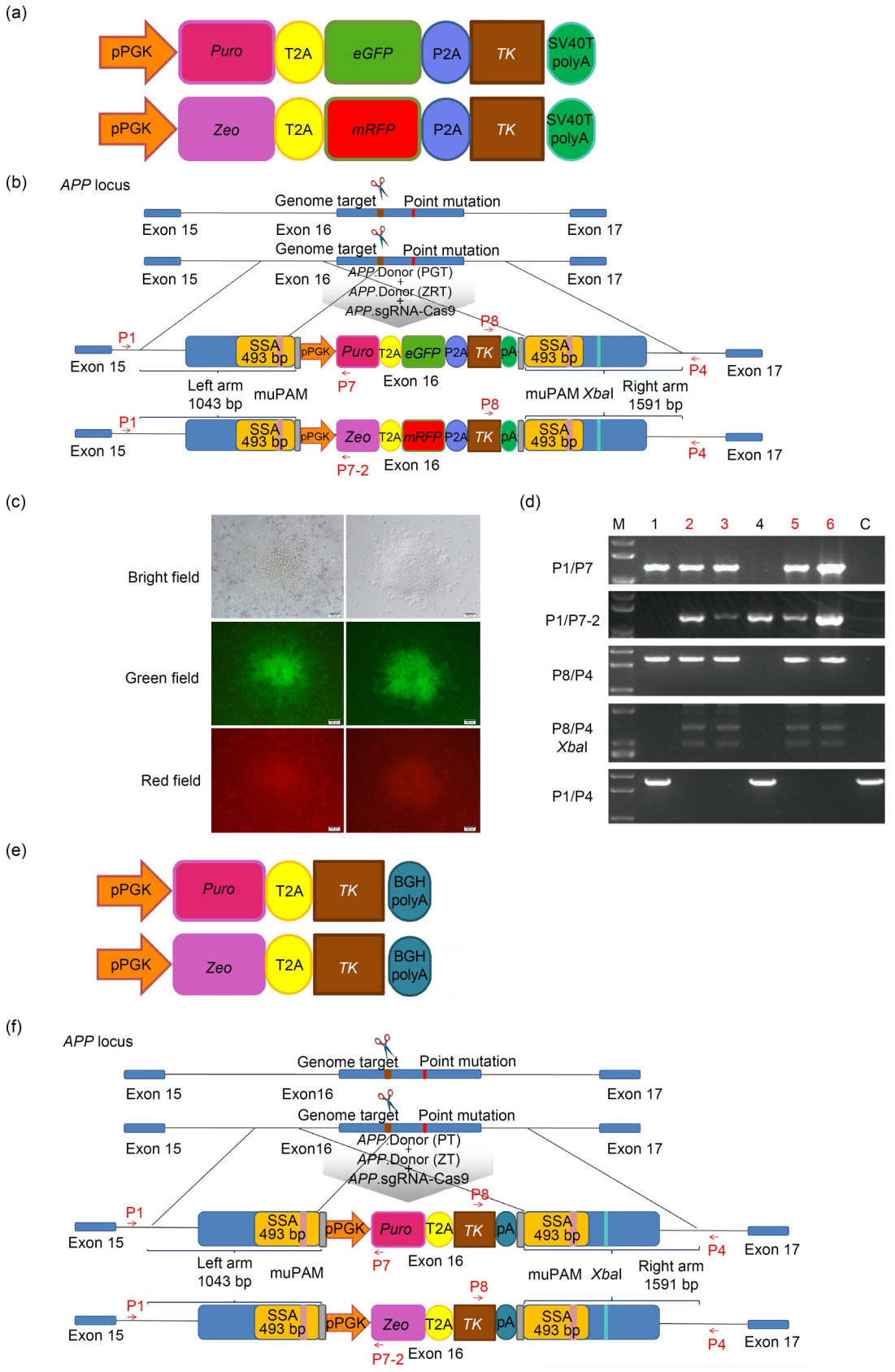
Fig. 2 TK-Puro-eGFP/TK-Zeo-mRFP (TPG/TZR)-dependent biallelic genome editing at the *PSEN1* locus. (a) Schematic diagram of the CRISPR/Cas9-mediated targeting and homology-directed repair dependent on biallelic editing at exon 3 of the *PSEN1* locus. Arrows indicate primers used for PCR analysis. Red arrows indicate primers used for PCR analysis. (b) Primer pairs (P5/P2, P3/P6 (TPG), P3-2/P6 (TZR), and P5/P6) were used to detect 5' and 3' recombinations and 5' and 3' biallelic integrations, respectively. *Bam*HI enzymatic digestion was applied to verify the introduction of the point mutation on P3/P6 (TPG) and P3-2/P6 (TZR) PCR fragments. C: control, a non-targeted HEK293T control sample; M: trans 2K plus DNA marker; Lanes 1 to 6: single colonies (red ones are positive colonies). (c) Sanger sequencing results of the PCR fragments amplified from the three homologous arms, confirming the desired point mutation at the *PSEN1* locus. *PSEN1*: presenilin 1; PCR: polymerase chain reaction; CRISPR: clustered regulatory interspaced short palindromic repeats; Cas9: CRISPR-associated protein 9 nuclease; sgRNA: single-guide RNA; pCAG: cytomegalovirus (CMV) immediate enhancer/β-actin (CAG) promoter; *TK*: thymidine kinase; pA: polyadenylic acid; pPGK: phosphoglycerate kinase promoter; *Puro*: puromycin resistance gene; *Zeo*: zeocin resistance gene; *eGFP*: enhanced green fluorescent protein; *mRFP*: monomeric red fluorescent protein; SSA: single-strand annealing; muPAM: mutated PAM.

arms and the one 3' homologous arm revealed the success of biallelic integration (Fig. S2a). Notably, our new PGT/ZRT system increased its biallelic targeting rate to 82.00% (41 of the 50 clones).

We finally constructed a third and even shorter pair of selection cassettes (PT/ZT) by further removing the fluorescent markers, which could also be used without the need for fluorescent microscopy or flow cytometry (Fig. 3e). *APP*.Donor (PT) and *APP*.Donor (ZT) (Figs. S1g and S1h) flanking the same homologous arm sequences of *APP* locus and were transfected together with *APP*.CRISPR/Cas9 into HEK293T cells (Fig. 3f). Without *eGFP* and *mRFP* fluorescent genes, we examined cell clones under a bright-field microscope (Fig. 3g). The detection primers located in *Puro*, *Zeo*, and *TK* gene sequences were not influenced by the deletion of the fluorescent protein gene located in the middle of the selection cassette; thus, we used the same primer pairs to detect biallelic targeted clones (Figs. 3b, 3f, and 3h). Through the same screening process of PCR amplification, restriction enzyme digestion, and Sanger sequencing, the PT/ZT system led to 70.00% positive clones (biallelic targeting in 35 out of 50 clones).

4 Discussion

Eukaryotic cells have evolved efficient systems to repair hazardous DSBs for cell survival. The


(a)
pPGK
Puro
T2A
eGFP
P2A
TK
SV40T polyA
pPGK
Zeo
T2A
mRFP
P2A
TK
SV40T polyA
(b)
APP locus
Genome target
Point mutation
Exon 15
Exon 16
Exon 17
APP.Donor (PGT)
+
APP.Donor (ZRT)
+
APP.sgRNA-Cas9
P1
P8
SSA 493 bp
pPGK
Puro
T2A
eGFP
P2A
TK
pA
SSA 493 bp
P4
P7
Left arm 1043 bp
muPAM
muPAM XbaI
Right arm 1591 bp
Zeo
mRFP
P7-2
(c)
Bright field
Green field
Red field
(d)
M 1 2 3 4 5 6 C
P1/P7
P1/P7-2
P8/P4
P8/P4 XbaI
P1/P4
(e)
pPGK
Puro
T2A
TK
BGH polyA
Zeo
(f)
APP locus
Exon16
APP.Donor (PT)
+
APP.Donor (ZT)
+
APP.sgRNA-Cas9

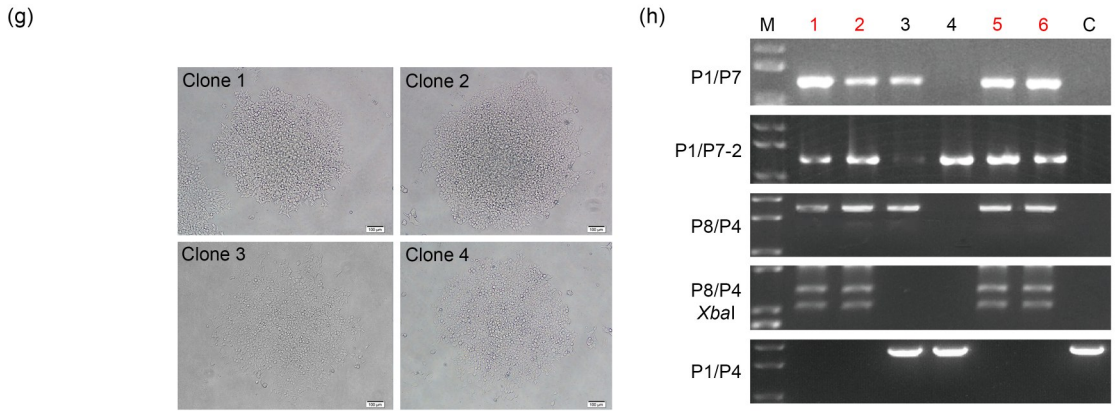


Fig. 3 Optimization of the selection cassettes and biallelic editing system. (a) Puro-eGFP-TK/Zeo-mRFP-TK (PGT/ZRT) selection cassettes. Shorter cassettes containing drug screening genes and fluorescent genes were constructed using the same PGK promoter and SV40T polyA. (b) PGT/ZRT-dependent biallelic genome editing at the *APP* locus. Targeting site and point mutation at the *APP* locus (top). Integration of the PGT/ZRT cassette and introduction of the point mutation according to homology-directed repair mediated by the CRISPR/Cas9 system (bottom). Red arrows indicate primers used for PCR analysis. (c) Representative fluorescent microscope photos of positively selected HEK293T clones after 10-d puromycin and zeocin screening. At this point, clones were picked and seeded into 24-well plates for expansion and analysis. Scale bar=100 μm. (d) Agarose gel images showing amplicons obtained by PCR amplification of the homologous arms using primer pairs P1/P7 (PGT), P1/P7-2 (ZRT), P8/P4, and P1/P4, respectively. The introduction of the point mutation on the P8/P4 amplicon was confirmed by *Xba*I restriction endonuclease digestion. C: control, a non-targeted HEK293T control sample; M: trans 2K plus DNA marker; Lanes 1 to 6: single colonies (red ones are positive colonies). (e) Puro-TK/Zeo-TK (PT/ZT) selection cassettes. These shorter selection cassettes were derived from the TPG plasmid; they retained the same promoter and polyA but only contained the drug selection gene; the fluorescence gene was removed. (f) PT/ZT-dependent biallelic genome editing at the *APP* locus. Targeting site and point mutation in exon 16 of the *APP* locus (upper part). Insertion of the PT/ZT cassette and introduction of the point mutation using the CRISPR/Cas9-mediated homology-directed repair; red arrows indicate primers used for PCR analysis. (g) Representative bright-field microscopy photos of the positively enriched HEK293T clones after 10-d puromycin and zeocin screening. Scale bar=100 μm. (h) Agarose gel images showing amplicons obtained using primer pairs P1/P7 (PT), P1/P7-2 (ZT), P8/P4, and P1/P4 to amplify the flanking homologous arms and ensure the biallelic editing separately. Digestion with *Xba*I enzyme was applied to verify the introduction of the correct targeted point mutation. C: HEK293T control without targeting; M: trans 2K plus DNA marker; Lanes 1 to 6: single colonies (red ones are positive colonies). PCR: polymerase chain reaction; CRISPR: clustered regularly interspaced short palindromic repeats; Cas9: CRISPR-associated protein 9 nuclease; sgRNA: single-guide RNA; pCAG: cytomegalovirus (CMV) immediate enhancer/β-actin (CAG) promoter; *TK*: thymidine kinase; pA: polyadenylic acid; pPGK: phosphoglycerate kinase promoter; *Puro*: puromycin resistance gene; *Zeo*: zeocin resistance gene; *eGFP*: enhanced green fluorescent protein; *mRFP*: monomeric red fluorescent protein; BGH: bovine growth hormone; *APP*: amyloid precursor protein; SSA: single-strand annealing; muPAM: mutated PAM.

non-homologous end joining (NHEJ) pathway, which introduces indels or alterations, can perform the DNA repair in most cell cycles (Bétermier et al., 2014; Zhao et al., 2017; Scully et al., 2019). In contrast, the HDR pathway, which induces high-fidelity repair, is only active in the S/G2 phase, and besides, its repair efficiency is highly dependent on the concentration of the donors (Saleh-Gohari and Helleday, 2004; Savic et al., 2018). DSBs can be introduced simultaneously on both alleles by the efficient CRISPR/Cas9 system. However, even in the presence of enough donors, in most cases, only one of the alleles is repaired by HDR; the ends of the other allele are joined by NHEJ (Xi et al., 2015; Paquet et al., 2016; Li et al., 2018). An ideal biallelic genome editing system should achieve high HDR efficiencies on both alleles in the same round of the repair process. Here, we designed and constructed an efficient, versatile, and programmable system for biallelic HDR-dependent gene editing that relies on dual screening donors. This TPG/TZR biallelic editing system achieved a positive rate of 52.94% and 54.90% at the *APP* and *PSEN1* loci, respectively, when targeted separately. A multicistronic structure using the self-cleaving 2A peptides allowed for even higher biallelic targeting efficiency of the *APP* locus: 82.00% for the PGT/ZRT system and 70.00% for the PT/ZT system (which was further edited to remove the fluorescent reporter genes). These results reveal that the shorter the inserted fragment, the higher the positive rate of biallelic targeted clones. Moreover,

the use of fluorescent selection genes appears to strengthen the screening of cell clones, but only to a certain extent. Of note, we used the TPG selection cassette to efficiently target *IGF2* and *RELA* loci in the pig kidney cell line (PK15), obtaining 30.0% and 53.3% efficiency, respectively (unpublished data). This further emphasizes the versatility of our selection systems. Interestingly, although the *TK* gene and single-strand annealing arm were introduced in the donor sequences to allow for potential deletion of the selection cassettes from the genome after the first round of HDR-induced repair, we failed to obtain colonies containing both marker-free alleles. This is possibly a result of introducing four DSBs into the genome simultaneously, which could be too stressful for chromosome repair. Further research and optimization of our gene editing system will be needed if deletion of the selection cassette is required after gene editing.

To put our novel strategy into perspective, we have previously tried to use dual surrogate reporter-integrated donors to screen biallelic gene-modified cells and demonstrated 34.09% efficiency at the C-C motif chemokine receptor 5 (*CCR5*) locus in HEK293T cells (Wu et al., 2017). Others have also reported low efficiency of biallelic gene editing; for instance, Eggenschwiler et al. (2016) achieved biallelic editing efficiency of only 8.70% at the *SERPINA* locus using a dual-fluorescence selection strategy, and ssODN has led to 1.2% to 8.3% accurate biallelic editing (Paquet et al., 2016). Relatively higher homozygously targeted efficiency (70%) has been reported by Koch et al. (2018) using CRISPR/Cas9 D10A nickase to introduce a GFP fusion protein that was smaller than 1 kb and flanked with relatively short homologous arms (500–800 bp). In this study, we were able to significantly improve those results, obtaining a high biallelic editing efficiency of up to 50%–82%, despite using over 2 kb inserts and 800–1600 bp homologous arms (Table 2).

Dosage compensation effects can minimize protein expression differences caused by genomic equivalence. In this case, both alleles need to be manipulated for a proper study of gene function or point mutation effects in disease models. When generating in vivo disease models, targeting of both alleles is of the essence. As an example, homozygous mutations of the *APP* (APP^{Swe}) or *PSEN1* ($PSEN1^{M146V}$) genes showed higher Alzheimer's disease-causing protein expression than heterozygous counterparts (Paquet et al., 2016). Furthermore, in regards to gene therapy, a homozygous genotype is often necessary to overcome genetic diseases (Evan, 2012; Sur et al., 2012), and heterozygous phenotypes are still not sufficient for restoring health (Ye et al., 2016; Frangoul et al., 2021). Our universal selection cassettes overcome the limitations imposed by the relatively low efficiency of HDR in biallelic genome editing. Although recent advances in base editing and prime editing techniques allow editing of the genome without a donor DNA supply, the traditional HDR pathway is still advantageous in biallelic editing with dual donors (Zhang, 2021).

5 Conclusions

We generated three pairs of selection cassettes to achieve high efficiency of biallelic editing at different loci, allowing the insertion of 2–5 kb fragments into the genome flanked with 0.8–1.6 kb homology arms for HDR. *APP* and *PSEN1* loci were chosen to be tested in HEK293T cells, and reached up to 82.00% positive efficiency. Our proof-of-concept studies strongly suggest that the three integrated selection systems we generated could be easily adapted to improve the targeting efficiency of virtually any other target locus. Regardless of the genome modification (knock-out or knock-in), homozygous cell lines not only provide more rigorous

Table 2 Summary of the developed biallelic editing systems

Biallelic system	Targeted locus	Drug resistance	Fluorescent reporter	Selection cassette (kb)	Homology arms (kb)	Ratio of positive clones	Biallelic targeting efficiency (%)
TPG/TZR	*APP*	Puromycin/zeocin	eGFP/mRFP	5.1/4.9	1.0/1.6	27/51	52.94
TPG/TZR	*PSEN1*	Puromycin/zeocin	eGFP/mRFP	5.1/4.9	0.8/1.0	28/51	54.90
PGT/ZRT	*APP*	Puromycin/zeocin	eGFP/mRFP	3.2/2.9	1.0/1.6	41/50	82.00
PT/ZT	*APP*	Puromycin/zeocin	None	2.4/2.2	1.0/1.6	35/50	70.00

TGP: TK-Puro-eGFP; TZR: TK-Zeo-mRFP; PGT: Puro-eGFP-TK; ZRT: Zeo-mRFP-TK; PT: Puro-TK; ZT: Zeo-TK; TK: thymidine kinase; Puro: puromycin resistance gene; eGFP: enhanced green fluorescent protein; Zeo: zeocin resistance gene; mRFP: monomeric red fluorescent protein.

isogenic controls but also exhibit more typical genome-dependent disease-associated phenotypes. Consequently, based on extremely high biallelic knock-in efficiency, our biallelic editing system offers a good choice for disease model generation.

Acknowledgments

This work was supported by the National Science and Technology Major Project of China (No. 2018ZX08010-09B), the Swedish Research Council (No. NE 2016-04458), and the Ragnar Söderberg Foundation (No. M21/17), Sweden. We would like to thank all the members of Professor ZHANG's lab for their excellent technical assistance and helpful discussions.

Author contributions

Xinyi LI and Zhiying ZHANG conceived and supervised the project. Xinyi LI, Bing SUN, Hongrun QIAN, and Jinrong MA conducted experiments and data analysis. Xinyi LI wrote the manuscript and edited it in collaboration with Magdalena PAOLINO. All authors have read and approved the final manuscript, and therefore, have full access to all the data in the study and take responsibility for the integrity and security of the data.

Compliance with ethics guidelines

Xinyi LI, Bing SUN, Hongrun QIAN, Jinrong MA, Magdalena PAOLINO, and Zhiying ZHANG declare that they have no conflict of interests.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Anzalone AV, Randolph PB, Davis JR, et al., 2019. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, 576(7785):149-157. https://doi.org/10.1038/s41586-019-1711-4

Bétermier M, Bertrand P, Lopez BS, 2014. Is non-homologous end-joining really an inherently error-prone process? *PLoS Genet*, 10(1):e1004086. https://doi.org/10.1371/journal.pgen.1004086

Bibikova M, Beumer K, Trautman JK, et al., 2003. Enhancing gene targeting with designed zinc finger nucleases. *Science*, 300(5620):764. https://doi.org/10.1126/science.1079512

Boch J, Scholze H, Schornack S, et al., 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, 326(5959):1509-1512. https://doi.org/10.1126/science.1178811

Chen XY, Janssen JM, Liu J, et al., 2017. In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. *Nat Commun*, 8:657. https://doi.org/10.1038/s41467-017-00687-1

Chu VT, Weber T, Wefers B, et al., 2015. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotechnol*, 33(5):543-548. https://doi.org/10.1038/nbt.3198

Codner GF, Mianné J, Caulder A, et al., 2018. Application of long single-stranded DNA donors in genome editing: generation and validation of mouse mutants. *BMC Biol*, 16:70. https://doi.org/10.1186/s12915-018-0530-7

Cong L, Ran FA, Cox D, et al., 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science*, 339(6121):819-823. https://doi.org/10.1126/science.1231143

Eggenschwiler R, Moslem M, Fráguas MS, et al., 2016. Improved bi-allelic modification of a transcriptionally silent locus in patient-derived iPSC by Cas9 nickase. *Sci Rep*, 6:38198. https://doi.org/10.1038/srep38198

Evan G, 2012. Taking a back door to target Myc. *Science*, 335(6066):293-294. https://doi.org/10.1126/science.1217819

Frangoul H, Altshuler D, Cappellini MD, et al., 2021. CRISPR-Cas9 gene editing for sickle cell disease and β-thalassemia. *N Engl J Med*, 384(3):252-260. https://doi.org/10.1056/NEJMoa2031054

Hsu PD, Lander ES, Zhang F, 2014. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, 157(6):1262-1278. https://doi.org/10.1016/j.cell.2014.05.010

Jinek M, Chylinski K, Fonfara I, et al., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337(6096):816-821. https://doi.org/10.1126/science.1225829

Koch B, Nijmeijer B, Kueblbeck M, et al., 2018. Generation and validation of homozygous fluorescent knock-in cells using CRISPR-Cas9 genome editing. *Nat Protoc*, 13(6):1465-1487. https://doi.org/10.1038/nprot.2018.042

Li C, Brant E, Budak H, et al., 2021. CRISPR/Cas: a Nobel Prize award-winning precise genome editing technology for gene therapy and crop improvement. *J Zhejiang Univ-Sci B (Biomed & Biotechnol)*, 22(4):253-284. https://doi.org/10.1631/jzus.B2100009

Li XY, Bai YC, Cheng XZ, et al., 2018. Efficient SSA-mediated precise genome editing using CRISPR/Cas9. *FEBS J*, 285(18):3362-3375. https://doi.org/10.1111/febs.14626

Mali P, Yang LH, Esvelt KM, et al., 2013. RNA-guided human genome engineering via Cas9. *Science*, 339(6121):823-826. https://doi.org/10.1126/science.1232033

Paquet D, Kwart D, Chen A, et al., 2016. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature*, 533(7601):125-129. https://doi.org/10.1038/nature17664

Ren CH, Xu K, Liu ZT, et al., 2015. Dual-reporter surrogate systems for efficient enrichment of genetically modified cells. *Cell Mol Life Sci*, 72(14):2763-2772. https://doi.org/10.1007/s00018-015-1874-6

Saleh-Gohari N, Helleday T, 2004. Conservative homologous recombination preferentially repairs DNA double-strand breaks in the S phase of the cell cycle in human cells. *Nucleic Acids Res*, 32(12):3683-3688. https://doi.org/10.1093/nar/gkh703

Savic N, Ringnalda FC, Lindsay H, et al., 2018. Covalent linkage of the DNA repair template to the CRISPR-Cas9 nuclease enhances homology-directed repair. *eLife*, 7: e33761. https://doi.org/10.7554/eLife.33761

Scully R, Panday A, Elango R, et al., 2019. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat Rev Mol Cell Biol*, 20(11):698-714. https://doi.org/10.1038/s41580-019-0152-0

Shao SM, Ren CH, Liu ZT, et al., 2017. Enhancing CRISPR/Cas9-mediated homology-directed repair in mammalian cells by expressing *Saccharomyces cerevisiae* Rad52. *Int J Biochem Cell Biol*, 92:43-52. https://doi.org/10.1016/j.biocel.2017.09.012

Sur IK, Hallikas O, Vähärautio A, et al., 2012. Mice lacking a *Myc* enhancer that includes human SNP rs6983267 are resistant to intestinal tumors. *Science*, 338(6112):1360-1363. https://doi.org/10.1126/science.1228606

Wang G, Yang LH, Grishin D, et al., 2017. Efficient, footprint-free human iPSC genome editing by consolidation of Cas9/CRISPR and piggyBac technologies. *Nat Protoc*, 12(1):88-103. https://doi.org/10.1038/nprot.2016.152

Wu Y, Xu K, Ren CH, et al., 2017. Enhanced CRISPR/Cas9-mediated biallelic genome targeting with dual surrogate reporter-integrated donors. *FEBS Lett*, 591(6):903-913. https://doi.org/10.1002/1873-3468.12599

Xi LH, Schmidt JC, Zaug AJ, et al., 2015. A novel two-step genome editing strategy with CRISPR-Cas9 provides new insights into telomerase action and *TERT* gene expression. *Genome Biol*, 16:231. https://doi.org/10.1186/s13059-015-0791-1

Yan NN, Sun YS, Fang YY, et al., 2020. A universal surrogate reporter for efficient enrichment of CRISPR/Cas9-mediated homology-directed repair in mammalian cells. *Mol Ther Nucleic Acids*, 19:775-789. https://doi.org/10.1016/j.omtn.2019.12.021

Yang LH, Guell M, Byrne S, et al., 2013. Optimization of scarless human stem cell genome editing. *Nucleic Acids Res*, 41(19):9049-9061. https://doi.org/10.1093/nar/gkt555

Ye L, Wang JM, Tan YT, et al., 2016. Genome editing using CRISPR-Cas9 to create the HPFH genotype in HSPCs: an approach for treating sickle cell disease and β-thalassemia. *Proc Natl Acad Sci USA*, 113(38):10661-10665. https://doi.org/10.1073/pnas.1612075113

Zhang BH, 2021. CRISPR/Cas gene therapy. *J Cell Physiol*, 236(4):2459-2481. https://doi.org/10.1002/jcp.30064

Zhang JP, Li XL, Li GH, et al., 2017. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol*, 18:35. https://doi.org/10.1186/s13059-017-1164-8

Zhao X, Wei CW, Li JJ, et al., 2017. Cell cycle-dependent control of homologous recombination. *Acta Biochim Biophys Sin*, 49(8):655-668. https://doi.org/10.1093/abbs/gmx055

Zuo EW, Sun YD, Wei W, et al., 2019. Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science*, 364(6437):289-292. https://doi.org/10.1126/science.aav9973

Supplementary information

Figs. S1 and S2