International Journal of Biological Macromolecules 243 (2023) 124926

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: www.elsevier.com/locate/ijbiomac

Efficient CRISPR/Cas9-mediated gene editing in mammalian cells by the novel selectable traffic light reporters

Ming Lyu [a], Yongsen Sun [b], Nana Yan [b], Qiang Chen [c], Xin Wang [a], Zehui Wei [a], Zhiying Zhang [a], Kun Xu [a,*]

[a] Key Laboratory of Animal Genetics, Breeding and Reproduction of Shaanxi Province, College of Animal Science and Technology, Northwest A&F University, Yangling 712100, China
[b] Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China
[c] Shaanxi Stem Cell Engineering and Technology Research Center, College of Veterinary Medicine, Northwest A&F University, Yangling 712100, China

ARTICLE INFO

Keywords:
CRISPR/Cas9
Traffic light reporter
Mammalian cells

ABSTRACT

CRISPR/Cas9 is a powerful tool for gene editing in various cell types and organisms. However, it is still challenging to screen genetically modified cells from an excess of unmodified cells. Our previous studies demonstrated that surrogate reporters can be used for efficient screening of genetically modified cells. Here, we developed two novel traffic light screening reporters, puromycin-mCherry-EGFP (PMG) based on single-strand annealing (SSA) and homology-directed repair (HDR), respectively, to measure the nuclease cleavage activity within transfected cells and to select genetically modified cells. We found that the two reporters could be self-repaired coupling the genome editing events driven by different CRISPR/Cas nucleases, resulting in a functional puromycin-resistance and EGFP selection cassette that can be afforded to screen genetically modified cells by puromycin selection or FACS enrichment. We further compared the novel reporters with different traditional reporters at several endogenous loci in different cell lines, for the enrichment efficiencies of genetically modified cells. The results indicated that the SSA-PMG reporter exhibited improvements in enriching gene knockout cells, while the HDR-PMG system was very useful in enriching knock-in cells. These results provide robust and efficient surrogate reporters for the enrichment of CRISPR/Cas9-mediated editing in mammalian cells, thereby advancing basic and applied research.

1. Introduction

Biological sciences have long sought better ways to edit the genetic code in cultured cells and in organisms to correct target genes through activation or deactivation (knock-in or knockout), in order to provide beneficial applications, from agricultural to biomedical. The rapid development of programmable gene editing tools has made it increasingly easier to broaden our knowledge of gene function and biological mechanisms. Meganuclease was first applied to genome editing in the 1990s, since then, zinc finger nucleases (ZFNs) [1], LAGLIDADG homing endonucleases (LHEs) [2], transcription activator-like effector nucleases (TALENs) [3], and clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated (CRISPR/Cas) nucleases have received extensive interest as efficient genome editing technologies [4–6]. Among these genome editing tools, the CRISPR/Cas system stands out as the most convenient, efficient and reliable tool. CRISPR is an adaptive immune system evolved in bacteria and archaea to protect from invasive bacteriophages or plasmids [7–9]. A a series of CRISPR/Cas systems (such as CRISPR/Cas9 and CRISPR/Cas12) have been adapted to edit mammalian genomes [10,11]. Currently, the most commonly used system is the *Streptococcus pyogenes* derived CRISPR/SpCas9, which consists of a single-guide RNA (sgRNA) and the SpCas9 nuclease [12,13]. The sgRNA directs SpCas9 nuclease to target the specific chromosomal DNA sequence by inducing sequence-specific double-strand break (DSB). Effective DSB induction and repair are essential for various genome modifications [14].

Cellular repair of targeted DNA DSBs is mainly mediated by three mechanisms: non-homologous end-joining (NHEJ), homology-directed repair (HDR), and single-strand annealing (SSA) [15]. NHEJ is an efficient but error-prone repair pathway that results in insertion/deletion

* Corresponding author.
E-mail address: xukunas@nwafu.edu.cn (K. Xu).

https://doi.org/10.1016/j.ijbiomac.2023.124926
Received 25 December 2022; Received in revised form 30 April 2023; Accepted 4 May 2023
Available online 20 May 2023

(indels) within the targeted locus. In contrast, HDR allows precise repair of a DSB by using an exogenous homologous donor DNA, which enables exact edits to target DNA sequence, including base substitutions, precise deletions and insertions of gene of interest in the presence of a donor flanked by homologous arms (HAs) [16,17]. SSA is a particular mechanism based on homologous recombination where a DSB repair occurs between direct repeats. The repeats recombine with each other, whereby one repeat and the sequence between the repeats are deleted [18,19]. Remarkably, the possible DSB repairment is determined by the competition between the various repair pathways, and the variability may be due to several factors such as the cell-cycle phase, DNA end complexity, transcriptional status around DSB sites, and the local chromatin structure [20].

Despite the success of CRISPR/Cas technologies for gene editing, their performance varies and is primarily determined by the quality and specificity of guide RNAs [9,13]. The ideal CRISPR/Cas tools should maintain the properties of compact Cas protein and broad targeting range with flexible protospacer adjacent motif (PAM) choice, with high activity and specificity. In most cases, CRISPR/Cas-mediated genome editing often occurs only in a limited fraction of transfected cells. Thus, a method that can be used for screening cells with high CRISPR/Cas nuclease activity would contribute to the enrichment of genetically modified cells. Currently, multiple tools and approaches have already been developed to facilitate the selection and enrichment of genetically modified cells. Among them, surrogate reporter systems have been proven effective in enriching genetically modified cells, as well as for rapid evaluation the DSB repair efficiency through different repair pathways [21–23]. Theoretically, an episomal surrogate reporter containing a target sequence might precisely reflect the given CRISPR/Cas nuclease activity and could be used for enrichment of the cells harboring corresponding edits in the genome [22,24]. Therefore, this method allows the identification and isolation of genetically modified cells when the editing efficiency is extremely low.

Previous studies have demonstrated that the surrogate reporters based on NHEJ and SSA repair mechanisms are promising for the enrichment of genetically modified cells [24,25]. Compared with transfection-positive screening strategies, the surrogate reporter-based strategies can achieve higher enrichment efficiency of genetically modified cells, taking advantage of nuclease-active screening by flow cytometric sorting, magnetic separation, or hygromycin selection [22,26]. However, the NHEJ- and SSA-based reporters are usually composed of segregated fluorescent gene expression cassettes as the transfection-marker and nuclease-active reporter, resulting in lower transfection and inconsistent fluorescence patterns, which limited seriously the application of the reporters. In addition, these surrogate reporters are prone to effective for enriching gene knockout cells. Usually, the HDR-based knock-in efficiency is pretty low (1–10 % of modified alleles) even in the enriched cells screened by the NHEJ- and SSA-based reporters.

A series of strategies have been tried to increase the HDR efficiency, including overexpressing key homologous recombination proteins (such as Rad51 and Rad52), inhibiting key proteins associated with NHEJ by RNAi (such as KU and LIG4) or using specific inhibitors (such as Scr7, and nocodazole), regulating cell cycle progression (such as nocodazole), and using different types of donor DNA templates (such as a conventional plasmid donor, linear donor, double-cut donor, and oligodeoxynucleotides) [27–30]. In addition, the efficiency of the HDR-mediated gene editing was supposed to be improved by the application of HDR-specific surrogate reporters.

In this study, we developed two puromycin-mCherry-EGFP (PMG) surrogate reporters based on SSA and HDR mechanisms, designated as SSA-PMG and HDR-PMG, respectively. The functional puromycin-resistance and EGFP selection cassette enabled both monitoring of nuclease activity in live cells by fluorescence and enrichment of genetically modified cells by puromycin selection and FACS. Firstly, the SSA-PMG and HDR-PMG surrogate reporters were successfully applied for the enrichment of CRISPR/Cas9-mediated genetically modified cells. Secondly, the different surrogate reporters were compared, and the SSA-PMG reporter exhibited the highest enrichment efficiency and sensitivity, which was prone to enrich gene knockout cells, while the HDR-PMG reporter was more suitable for the enrichment of knock-in cells. What's more, the HDR-PMG enrichment efficiency was further improved by collaborative usage of different donor forms, small molecules and other factors. In addition, the two surrogate reporters were adapted to other CRISPR/Cas systems (SlugCas9 [31] and AsCas12a [32]), and were also applied for efficient gene editing in primary mammalian cells. Taken together, the novel SSA-PMG and HDR-PMG screening traffic light reporters can provide rapid, simple and efficient enrichment of genetically modified cells, which will contribute to the practical application of mammalian gene editing technologies.

2. Experimental procedures

2.1. Ethics statements

All the procedures of animal experimentation in this study strictly followed the protocol approved by the Institutional Animal Care and Use Committee (IACUC) of Northwest Agricultural and Forestry University (NWAFU) (permit number: 15–516, date:9–13–2015).

2.2. Plasmids construction

All polymerase chain reactions (PCR) for molecular cloning was performed using PrimeSTAR Max DNA Polymerase (R045Q, Takara Bio Inc., Otsu, Japan), unless otherwise indicated. The restriction enzyme digestion and T4 DNA ligation reactions were performed according to the manufacturer's instructions (Takara). PCR primers and DNA oligonucleotides were synthesized with Sangon Biotech (Shanghai, China). All PCR products and intermediate plasmid products were confirmed *via* Sanger sequencing in both directions (Sangon). The CRISPR/Cas targets used in this study are listed in Table 1.

2.3. Construction of the SSA-PMG surrogate reporter

The SSA-PMG vector was constructed by a modular cloning strategy. To construct the SSA-PMG plasmid, a DNA sequence encoding mCherry

Table 1
CRISPR/Cas targets used in this study.

Locus	Species	Nuclease type	Target sequences	PAM
AAVS1	Human	CRISPR/Cas9	CTGTCCCCTCCACCCCACAG	TGG
CCR5	Human	CRISPR/Cas9	CACACTTGTCACCACCCCAA	AGG
EMX1	Human	CRISPR/Cas9	GAGTCCGAGCAGAAGAAGAA	GGG
NUDT5	Human	CRISPR/Cas9	GTGAAGTGTTCTCTGCAGCA	CGG
VEGFA	Human	CRISPR/Cas9	CTCGGCCACCACAGGGAAGC	TGG
VEGFA	Human	CRISPR/SlugCas9	GCTCGGCCACCACAGGGAAGC	TGGG
VEGFA	Human	CRISPR/AsCas12a	CTAGGAATATTGAAGGGGGCAGG	TTTG
Rosa26	Mouse	CRISPR/Cas9	ACTCCAGTCTTTCTAGAAGA	TGG
IGF2	Porcine	CRISPR/Cas9	TTCGCCTAGGCTCGCAGCGC	GGG
INSIG1	Bovine	CRISPR/Cas9	GGTGATGCCAGCTGTTGACG	TGG
MSTN	Goat	CRISPR/Cas9	CGATGACTACCACGTTACGA	CGG
MyoG	Goat	CRISPR/Cas9	GGAACCTCACTTCTATGACG	GGG

was PCR-amplified from pAAV-minCMV-mCherry (Addgene, # 27970), and cloned into an SSA-based reporter plasmid in our lab (Addgene, #85932) between *BamH* I and *Not* I, generating PuroL(1–305)-mCherry-PuroR(105–597)-T2A-EGFP. The fragment was then inserted through the *Hind* III/*Xba* I restriction site into the pcDNA3.1(+) vector with the cytomegalovirus (CMV) promoter. At this time, the parental SSA-PMG vector was constructed with only the functional fluorescent *mCherry* gene as a transfection-marker.

To construct the target specific SSA-PMG vectors, the *mCherry* gene sequence flanked with CRISPR/Cas targets was amplified. Next, this fragment was inserted through the *BamH* I and *Not* I sites into the SSA-PMG vector generating the gene of interest (GOI)-SSA-PMG (pGOI-SSA-PMG). When the CRISPR/Cas nuclease functioned to recognize and cleave the targets within the surrogate reporter, the SSA-mediated repair would result in the correction of the open reading frame (ORF) for the puromycin-resistance and EGFP selection cassette. Primer sequences are provided in Supplementary Table S1.

2.4. Construction of the HDR-PMG surrogate reporter

For the HDR-PMG surrogate reporter, two truncated puromycin-resistance gene coding sequences: PuroL(1–225) and PuroR(330–597), were separated using the mCherry fragment. In parallel, a truncated (ATG removed) and inverted $Puro^R$ fragment (ΔPuro) was integrated into the pcDNA3.1(+) plasmid backbone to yield the parental HDR-PMG plasmid.

Similar as above, the *mCherry* gene sequence flanked with CRISPR/Cas targets was amplified and inserted into the HDR-PMG plasmid to generate the GOI-HDR-PMG vector. When the CRISPR/Cas nuclease cleaved the target sequence, the HDR-mediated DSB repair would generate a correct ORF and delete the intermediate sequences of the *mCherry* gene, resulting in functional expression of the puromycin-resistance and EGFP selection cassette. Primer sequences are provided in Supplementary Table S1.

2.5. Construction of the sgRNA vectors and donor plasmids

For the construction of pGOI-sgRNA/Cas9, synthetic reverse complementary oligonucleotides (Sangon) were annealed to generate double-stranded oligonucleotides and then cloned into pX330-U6-Chimeric_BB-CBh-hSpCas9 (Addgene, #42230). The primer sequences are listed in Supplementary Table S1.

For the construction of the conventional plasmid donor, the donor DNA fragment with homologous arms (approximately 1600–2000 bp) flanking the CRISPR/Cas target was obtained by overlap PCR with the genomic DNA as the template, and was integrated into pXL-BACII (Addgene, #170519) to generate the corresponding donor vector (pD-GOI). The PAM of the target within the donor vector was replaced by a restriction endonuclease recognition site for the future digestion assay to measure HDR efficiency.

For the double-cut donor plasmid (pD-GOI/sgRNA), the major difference from the pD-GOI plasmid donor was the addition of CRISPR/Cas targets flanking the homologous arms [33]. Primer sequences are shown in Supplementary Table S2. Additionally, single-strand oligodeoxynucleotide donor (ssODN) templates of different lengths were synthesized and compared for the HDR editing efficiency (Table S3).

2.6. Cell culture, transfection, and treatment

HEK293T cells (CRL-11268, ATCC, Rockville, MD, USA), HepG2 cells (ATCC, HB-8065), Hela cells (ATCC, CCL-2), and porcine embryonic fibroblasts (pEFs) were maintained in Dulbecco's modified Eagle medium (C11995500BT, Invitrogen Corp., Waltham, MA, USA) supplemented with 10 % fetal bovine serum (FBS, Gibco, Grand Island, NY, USA). Feeder-free mouse embryonic stem cells (mESCs) were cultured on 0.1 % gelatin-coated plates with DMEM supplemented with 15 % FBS (Gibco, Grand Island, NY, USA), 1000 units/mL recombinant leukemia inhibitory factor (LIF), 0.1 mM 2-mercaptoethanol, 2 mM L-glutamine, 0.1 mM MEM non-essential amino acids (NEAA), and 1 % nucleoside mix (100× stock, Sigma Aldrich, MO, USA). Bovine mammary epithelial cells (BMECs) were grown in DMEM/F12 and 10 % FBS, with a variety of cytokines added (such as 5 μg/mL bovine insulin). Primary goat myoblasts were cultured in growth medium containing DMEM/F12 basic medium (Gibco, Grand Island, NY, USA), and 20 % FBS (Gibco, Grand Island, NY, USA). All media were supplemented with antibiotics (100 units /mL penicillin and streptomycin) (15140–122, Invitrogen Corp., Waltham, MA, USA). All type of cells used in this study were maintained under standard culture conditions (37 °C, 5 % CO_2 incubator). The culture medium was changed every two days.

For transfection, the cells were seeded into 12-well plates 24 h prior to transfection and were transfected with Lipofectamine™ 3000 reagent (L3000015, Invitrogen Corp., Waltham, MA, USA) according to the supplier's protocol. All experiments with small molecule compounds were performed as we previously described [23]. For Scr7 treatment, HEK293T cells were incubated with Scr7 (1 mM, Sigma) 24 h after transfection until cells were collected. For nocodazole treatment, nocodazole (100 ng/mL, Sigma) was added 12 h after transfection for 24 h and then released.

2.7. Cell viability assays

Cell viability was evaluated using a Cell Counting Kit-8 (CCK-8) assay (Dojindo Molecular Technologies, Inc., Kumamoto, Japan) following the manufacturer's instructions. HEK293T cells were seeded at a density of 1×10^3 cells per well into a 96-well plate using a Countstar Automated Cell Counter (Countstar BioTech, Shanghai, China). Three independent biological replicates for each treatment were conducted. Three days after transfection, cells were selected by puromycin, and the timepoint was defined as 0 h. At the time points 0 h, 24 h, 48 h, 72 h, and 96 h, the cells were switched to the medium with 10 % CCK-8, and the absorbance was measured at 450 nm using a SYNERGY/H1 microplate reader (BioTek, Winooski, VT, USA).

2.8. Protein extraction and Western blot analysis

Cells were harvested 3 days after the transfection and subjected to protein extraction. The cell samples were lysed in RIPA buffer (Solarbio, Beijing, China) according to the manufacturer's instructions. The protein samples were separated by 12 % sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) followed by transfer onto a polyvinylidene fluoride (PVDF) membrane (Millipore, Burlington, MA, USA). The membrane was blocked with 5 % defatted milk (BD, Franklin Lakes, NJ, USA) in TBST (tris-buffered saline with 0.05 % Tween 20) buffer for 1 h at room temperature and then incubated overnight at 4 °C with antibodies against puromycin (1:1000 dilution, Thermo Fisher, Massachusetts, USA) and β-tubulin (1:2000 dilution, CWBIO, Beijing, China). The following day, the membrane was washed three times with TBST for 10 min each time. The membrane was then incubated with HRP-conjugated secondary antibodies (anti-mouse IgG or anti-rabbit IgG) (1:1000 dilution, CWBIO) for 1 h at room temperature. The chromogenic reaction was performed using an enhanced chemiluminescence (ECL) Western blot substrate (Advansta, Menlo Park, CA, USA) and was detected using a Sage Capture TM system (BioTek, Winooski, VT, USA).

2.9. Flow cytometry

The cells were cultured in 6-well plates at a density of 5×10^5 cells per well. Three days after transfection, the cells were washed with phosphate-buffered saline (PBS) and subsequently harvested. Next, the cells were resuspended in 1 mL of PBS and the self-repair efficiency of different surrogate reporter systems was accessed by counting the

percentage of EGFP positive cells by the BD FACSAria™ III flow cytometry system (Franklin Lakes, NJ, USA), and the data was analyzed using CytExpert software. EGFP-positive cells were subsequently harvested, while the remaining population was used as control cells. Genomic DNA was isolated using a TIANamp Genomic DNA Kit (Tiangen Biotech Co. Ltd., Beijing, China) and was subjected to subsequent experiments.

2.10. Detection of the knockout efficiency in genomic level

The T7E1 assay was performed as previously described [34]. The genomic region containing the intended target site was amplified by nested PCR using site-specific primers (Table S4). The PCR products (300 ng) were then denatured and reannealed to allow the formation of a DNA heteroduplex, which was treated with T7 endonuclease 1 (New England Biolabs, Hitchin, UK) for 20 min at 37 °C and analyzed by 2 % DNA gel electrophoresis. Mutation frequencies were calculated as previously described based on the band intensities using ImageJ software (https://mirror.imagej.net/) and the following equation: Mutation frequency (%) $=100 \times [1 - (1 - \text{fraction cleaved})^{1/2}]$ [34,35]. In addition, PCR-amplified fragments were also inserted into the pMD-19 T (6013, Takara) vector by T-A cloning and 15 or more clones were sequenced.

2.11. Analysis of HDR-mediated knock-in efficiency by restriction digestion

Digestion assays were performed as previously described [23]. Briefly, the target regions of different genomic loci were amplified by PCR using corresponding detection primers (Table S5) and purified. Then, the PCR products (300 ng) were digested with corresponding restriction endonucleases. The relative HDR frequency was quantified by the restriction enzyme digestion of PCR-amplified target genes: fraction cleaved = (sum of intensities of cleaved bands)/(sum of intensities of cleaved and uncut bands) [23]. On the other hand, PCR products were cloned into the pMD-19 T and >15 independent clones were sequenced from each sample.

2.12. Deep sequencing analysis

For further detection of the knockout efficiency, the genomic region flanking the CRISPR/Cas target site for each gene (~280 bp) was amplified by PCR with specific barcode primers. For the precise knock-in efficiency, the genomic region around the CRISPR/Cas target site for each gene was amplified by a two-step PCR method. Here, the PCR products obtained in 2.11 before the digestion assays were used as the template for the second round of PCR with primers flanking the barcode, as we previously described [23]. Amplicons were sequenced on an Illumina PE150 (Tsingke Biotechnology Co., Ltd. Wuhan, China). The sequencing data were analyzed with CRISPResso2 (http://www.crispresso.rocks).

2.13. Statistical analysis

All statistics were analyzed using GraphPad Prism software 8.0 and the data were indicated as the "means ± SEM". Statistical significance was examined by the unpaired Student's *t*-test for two-group comparisons or a one-way ANOVA for more than two groups. The data were considered significant and highly significant when the corresponding *P* value was <0.05 (*) and 0.01 (**).

3. Results

3.1. Generation and validation of the novel screening traffic light reporters, SSA-PMG and HDR-PMG

Based on our previous studies, we developed the SSA-PMG reporter to screen and enrich genomically modified cells with knockout events. To increase the sensitivity and efficiency of the surrogate reporter that we previously developed (Fig. S1) [34], the reporter construct was modified with dual-reporter genes (the puromycin- resistance and EGFP genes) and red fluorescent marker gene were assembled in a single cassette driven by the CMV promoter (Fig. 1A). In the SSA-PMG reporter construct, the coding sequence of the *mCherry* gene, flanked by two CRISPR/Cas9 nuclease targets, was inserted between two truncated puromycin-resistance gene sequences: PuroL(1–305 bp) and PuroR (105–597 bp) with a 200 bp direct repeat. The *EGFP* gene could not be expressed because there was a termination codon in front of the PuroR (105–597 bp) sequence (Fig. 1A). Theoretically, only functional mCherry fluorescent protein can be expressed by the CMV promoter before CRISPR/Cas9 targeting, which was intended to serve as a transfection marker and will be deleted after CRISPR/Cas9 targeting. The puromycin-resistance and EGFP fluorescence dual-reporter genes would not be expressed until the construct targeted by CRISPR/Cas9 nuclease and repaired by the SSA mechanism.

In addition, we constructed the HDR-PMG reporter based on the HDR-mediated DSB repair mechanism. The HDR-PMG reporter was designed similarly to the SSA-PMG reporter construct, except for different lengths of truncated puromycin-resistance gene sequences (PuroL,1–225 bp, and PuroR, 330–597 bp, without direct repeat) and an HDR donor template (ΔPuro, the puromycin-resistance gene sequence without an ATG and promoter) (Fig. 1B). As expected, HDR-mediated repair of a potential DSB in the target region will reconstitute the puromycin-resistant expression and green fluorescence. In principle, the puromycin-resistance and EGFP fluorescent dual-reporter genes could only be repaired by the HDR mechanism when the construct was targeted by the CRISPR/Cas9 nuclease.

To verify the SSA-PMG and HDR-PMG systems, a target sequence in the human *AAVS1* site (h*AAVS1*) was chosen to construct the SSA-PMG and HDR-PMG reporters, and the corresponding CRISPR/Cas9 targeting vector phAAVS1-sgRNA/Cas9 was constructed as well. The SSA-PMG and HDR-PMG reporters together with the targeting vector phAAVS1-sgRNA/Cas9 were used to co-transfect HEK293T cells. Cells co-transfected with the pPuro-T2A-EGFP plasmid (Fig. S2A) and phAAVS1-sgRNA/Cas9 were used as the positive control. Increasing numbers of EGFP-positive cells emerged following transfection in both the control and experimental groups, and decreasing mCherry fluorescence was observed in the experimental groups, indicating that the SSA-PMG and HDR-PMG systems functioned as expected (Fig. S2B-S2D). Additional negative non-targeting controls with co-transfection of pNC-sgRNA/Cas9 (Cas9 expression vector without sgRNA) and the PMG reporters further demonstrated that the EGFP reporter gene could not be restored without CRISPR/Cas9 targeting (Fig. 1C). Cells from different transfection groups were then subjected to a transient puromycin selection assay. The results showed that cells could only survive in the experiment and positive control groups, indicating that the puromycin-resistance gene had been repaired in the SSA-PMG and HDR-PMG reporters (Fig. 1D).

Taking into consideration the puromycin selection time and cell viability, we next examined the cell viability by CCK-8 assays. The results showed that the cell viability decreased as the puromycin treatment time increased, particularly in the puromycin-resistance gene non-expression group (Fig. 1E). Western blot analysis further demonstrated that the successful expression of the puromycin resistance protein in the CRISPR/Cas9-targeted surrogate reporter groups (Fig. 1F–1H). Overall, the results described above demonstrated that the SSA-PMG and HDR-PMG surrogate reporters can be activated by CRISPR/Cas9 to generate functional reporter genes, which will be applied for the enrichment of genetically modified cells.

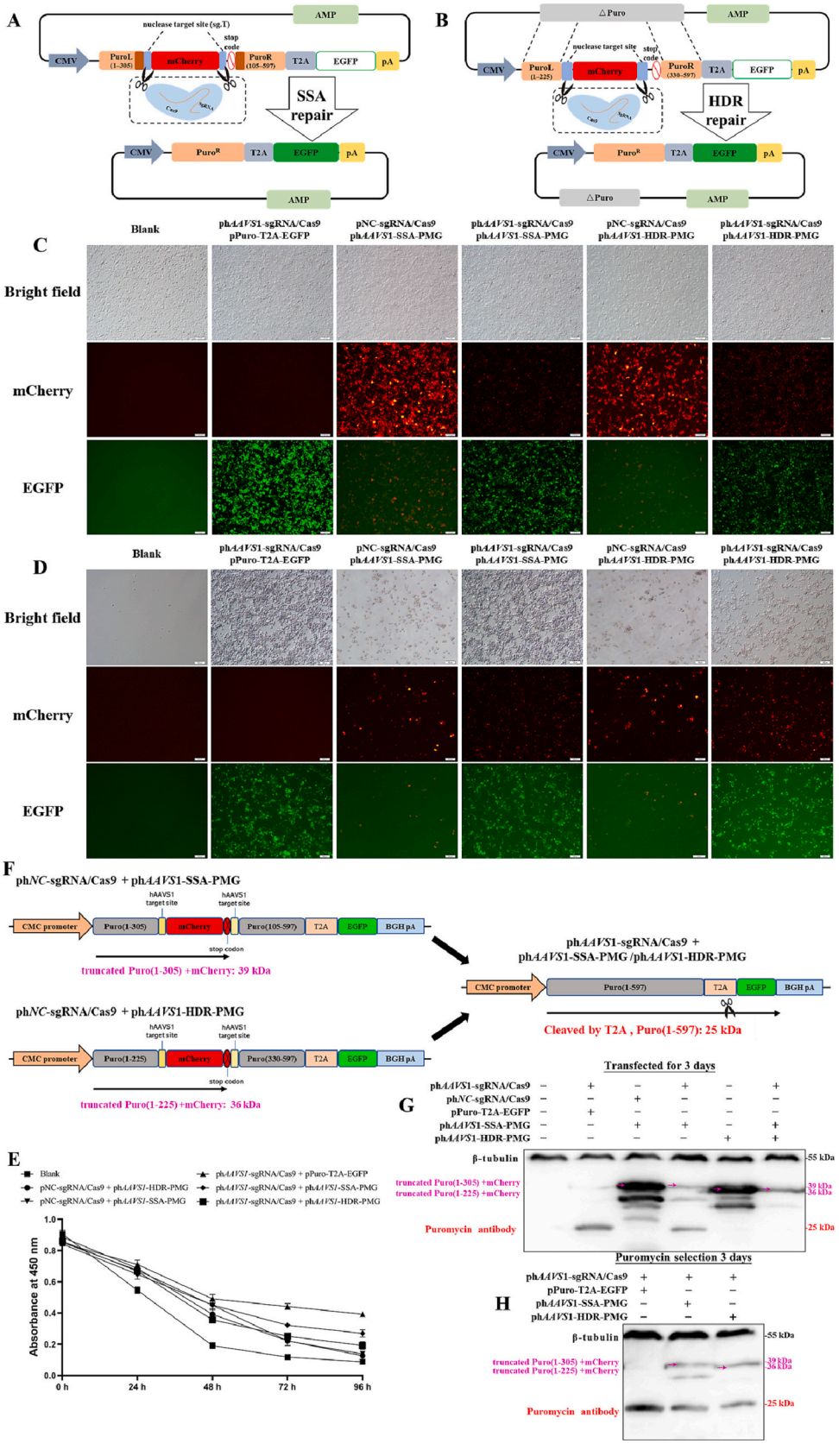


Fig. 1. Function principle and feasibility proof of the SSA-PMG and HDR-PMG surrogate reporters.
(A) Schematic representation of the SSA-PMG reporter. *mCherry* is constitutively expressed by the CMV promoter. The target-disrupted puromycin-resistance gene (Puro^R) and *EGFP* fluorescence gene are fused by T2A as dual reporter, whereas Puro^R and EGFP are not expressed without CRISPR/Cas9 targeting because the puromycin-resistance cassette is interrupted by the *mCherry* sequence with sgRNA targets at both ends, flanked with direct repeats as SSA arms, and there is a stop codon before EGFP. When intent CRISPR/Cas9 is introduced, DSBs can be generated and repaired by SSA, resulting in the deletion of one repeat and the intervening *mCherry* marker gene, correcting the ORF for puromycin-resistance and EGFP reporter gene cassette. (B) Schematic representation of the HDR-PMG reporter. The reporter construct contains a CMV promoter-driven incomplete puromycin-resistance sequence that deleted 104 bp, with the *mCherry* sequence flanked by sgRNA targets inserted between the two truncated puromycin coding sequences. Additionally, a puromycin-resistance sequence without ATG (ΔPuro) was also cloned serving as the HDR template. When the sgRNA associates with Cas9 and directs the nuclease to the targets in the reporter, functional puromycin-resistance and EGFP reporter cassette can only be generated by the HDR-based DSB repairing. (C) Visualization of mCherry and EGFP expression by fluorescence microscopy. HEK293T cells were co-transfected with ph*AAVS*1-sgRNA/Cas9 expression plasmid and corresponding surrogate reporter plasmids. As the control group, the pNC-sgRNA/Cas9 (Cas9 expression vector without sgRNA) was used. Cells were observed and photographed under a fluorescence microscope three days after transfection. (D) Transfected cells were observed and photographed after maintained continuously with puromycin for 3 days. (E) CCK-8 assay for the viability of cells transfected with different reporters after puromycin selection. (F) Schematic representation of puromycin-resistance protein expression before and after SSA-PMG and HDR-PMG repairing. (G) Western blot assay for the detection of puromycin-resistance protein expressed in cells three days after transfection. (H) Western blot detection of the puromycin-resistance protein in enriched cells after puromycin selection. Scale bar = 100 μm.

3.2. Enrichment of genetically modified cells with SSA-PMG and HDR-PMG reporters

To further demonstrate that the SSA-PMG and HDR-PMG surrogate reporters could be used to enrich genetically modified cells, we next examined ph*AAVS*1-SSA-PMG and ph*AAVS*1-HDR-PMG for the editing of the h*AAVS*1 locus in HEK293T cells (Fig. 2A). Three days after transfection, the genetically modified cells were enriched using the two different screening assays (FACS and puromycin selection). FACS results showed that 32.94 % of cells were $EGFP^+$ in the SSA-PMG-h*AAVS*1 group (transfected with ph*AAVS*1-sgRNA/Cas9 and ph*AAVS*1-SSA-PMG), and 37.97 % of cells were $mCherry^+$ in the SSA-PMG-h*AAVS*1 control group (transfected with ph*NC*-sgRNA/Cas9 and ph*AAVS*1-SSA-PMG). In addition, only 18.58 % of cell were $EGFP^+$ in the HDR-PMG-h*AAVS*1 group (transfected with ph*AAVS*1-sgRNA/Cas9, ph*AAVS*1-HDR-PMG, and pD-h*AAVS*1), compared to 35.22 % of cells that were $mCherry^+$ in the HDR-PMG-h*AAVS*1 control group (transfected with ph*NC*-sgRNA/Cas9, ph*AAVS*1-HDR-PMG, and pD-h*AAVS*1) (Fig. 2B). As expected, HEK293T cells transfected with either ph*AAVS*1-sgRNA/Cas9, ph*AAVS1*-SSA-PMG, ph*AAVS*1-HDR-PMG, or pD-h*AAVS1* alone did not result in any $EGFP^+$ cells, and the cells in all the control groups died after puromycin selection. On the other hand, the association of the sgRNA/Cas9 plasmid with the surrogate reporters resulted in $EGFP^+$ and $Puro^{R+}$ cells (Fig. 2C).

Genomic DNA was extracted from the positive cells ($EGFP^+$ cells or $Puro^{R+}$ cells) and was subjected to PCR genotyping. The T7E1 assay results of the SSA-PMG reporter groups demonstrated significant increase of mutation frequency in cells enriched by FACS and puromycin selection, which was 5.1- and 6.2-fold of that in unsorted cells, respectively (Fig. 2D), indicating that the SSA-PMG reporter enables efficient enrichment of mutant cells. The sequencing results further confirmed the enrichment by revealing 40.0 % and 60.0 % mutation frequency in the FACS-selected and puromycin-resistant cells, respectively, which was 6.4- and 9.6-fold of that in unsorted cells, (Fig. 2E). The result of the restriction enzyme digestion assay (Fig. 2F) of the HDR-PMG reporter groups showed 13.8 % and 20.2 % HDR-based knock-in mutation efficiency in cells after flow sorting and puromycin screening, respectively (Fig. 2G), while the HDR events were nearly undetectable in unsorted cells. This result indicated that our HDR-PMG reporter can be used specifically for the enrichment of HDR-based genetically modified cells.

3.3. Comparison of the SSA-PMG and HDR-PMG surrogate reporters in self-repairing efficiency in different gene loci and cell types

The self-repairing efficiency of a surrogate reporter after targeted by the CRISPR/Cas nuclease is critical for the enrichment of nuclease-active cells. Therefore, we further tested the applicability of different surrogate reporters for different gene loci and cell types. Three gene loci (*CCR5*, *EMX1*, and *NUDT5*) in HEK293T cells were selected, which are commonly used to evaluate the efficiency of a gene editing system.

Firstly, the self-repairing efficiency of the SSA-PMG reporter was compared with the NHEJ-RPG and SSA-RPG surrogate reporters (Fig. S1), which we have previously reported [34]. HEK293T cells were transfected with the CRISPR/Cas9 expression plasmids (ph*CCR*5-sgRNA/Cas9, ph*EMX*1-sgRNA/Cas9 or ph*NUDT*5-sgRNA/Cas9) and the corresponding surrogate reporter vectors (pGOI-NHEJ-RPG, pGOI-SSA-RPG or pGOI-SSA-PMG), and three days after transfection the cells were collected for flow cytometry counting analysis. The results demonstrated that the SSA-PMG group possessed the highest percentages of EGFP-positive cells, which was 1.22- to 1.23-fold of that in the NHEJ-RPG groups and 1.08- to 1.18-fold of that in the SSA-RPG groups (Fig. 3A–C). In addition, similar results were observed using two additional cancer cell lines (HepG2 and Hela) (Fig. 3D–E and Fig. S3), indicating that the novel SSA-PMG reporter enables efficient recapitulation of CRISPR/Cas9 nuclease activity in different cell types and is more sensitive than the previously developed reporters.

We next compared the sensitivities of the SSA-PMG and HDR-PMG surrogate reporters. As predicted, the percentage of EGFP-positive cells in the HDR-PMG group was much lower than that of the SSA-PMG group (Fig. 3F–H). We obtained similar results in HepG2 and Hela cell lines (Fig. 3I–J and Fig. S4). This implied a direct relationship between the sensitivity of the surrogate reporters and the repairing efficiency of the mechanisms applied. We therefore reasoned that surrogate reporters based on different repair mechanisms could be capable for the enrichment of different types of genetically modified cells.

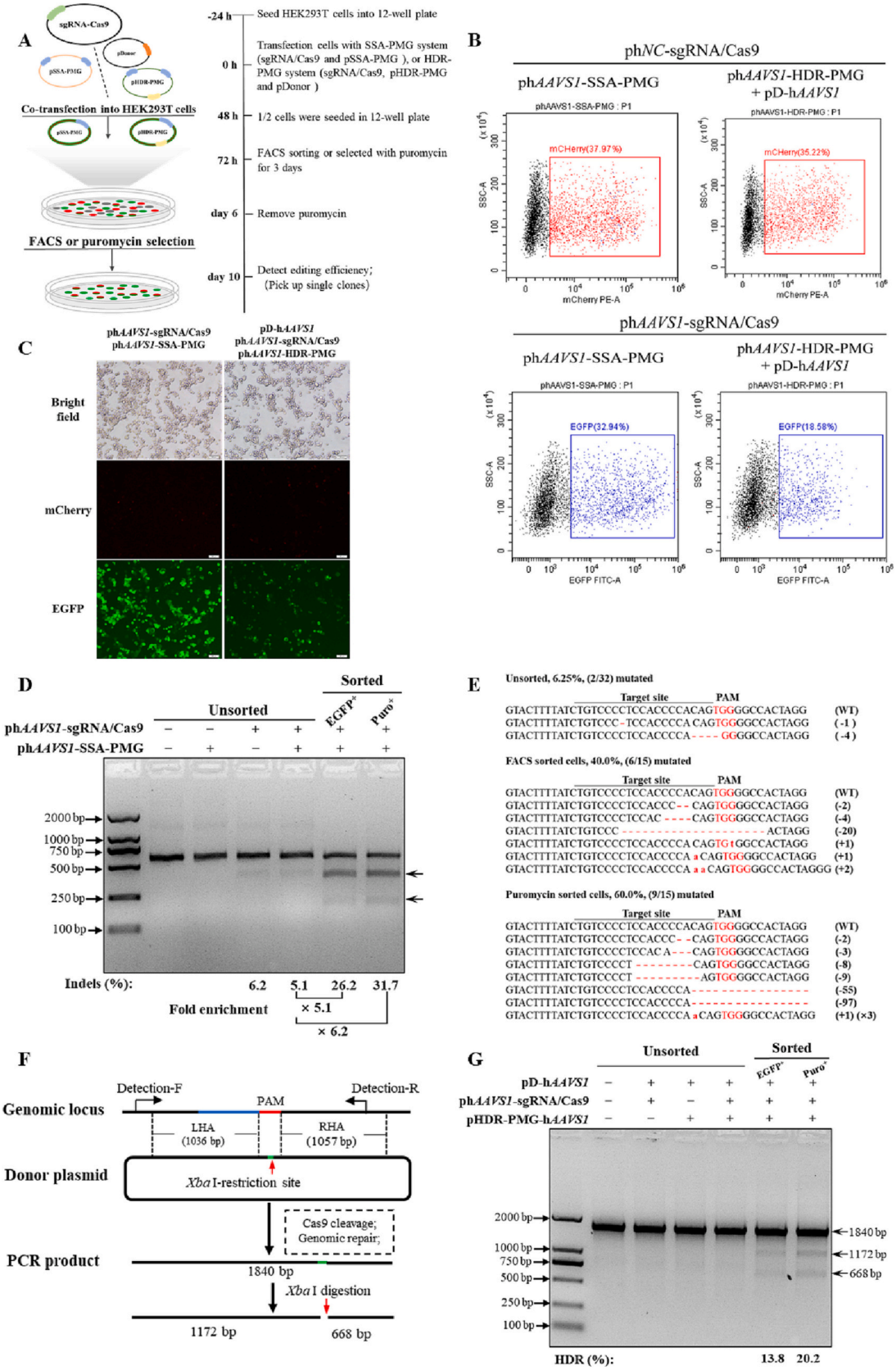
3.4. Detection of CRISPR/Cas9-mediated knockout efficiency in cells enriched by the SSA-PMG surrogate reporter

To further verify the superiority of the SSA-PMG surrogate reporter for the enrichment of genetically modified cells, we tested the CRISPR/Cas9-mediated knockout efficiency in cells enriched with different surrogate reporters using the two different screening methods: FACS and puromycin selection. HEK293T cells were co-transfected with the sgRNA/Cas9 expression vectors and corresponding reporter plasmids (pPuro-T2A-EGFP, pGOI-NHEJ-RPG, pGOI-SSA-RPG or pGOI-SSA-PMG). Three days after transfection, positive cells were screened for green fluorescence or puromycin resistance.

As shown above, the SSA-PMG reporter exhibited higher self-repairing efficiency and was more sensitive, suggesting that the SSA-PMG reporter may be more suitable for enriching the nuclease-active cells, which will contribute to the enrichment of genetically modified cells. T7E1 assays revealed that the mutation frequencies at the h*CCR*5, h*EMX*1 and h*NUDT*5 genomic loci in FACS-selected cells were 28.8 %, 17.4 % and 18.4 %, respectively, which were 4.5-,11.6- and 20.4-fold of those in unsorted cells (6.4 %, 1.5 % and 0.9 %, respectively) (Fig. 4A and Fig. S5A—5B). The enrichment effect was also confirmed in cells enriched by puromycin selection, and we detected that the mutation frequencies at the three genomic loci were 43.1 %, 24.7 % and 23.1 %, respectively, and were 6.7-, 16.5- and 25.7-fold of those in unsorted cells (Fig. 4A and Fig. S5A—5B). Furthermore, sequencing of the genomic DNA around the target site showed that the mutation frequencies of FACS or puromycin-selected cells were significantly higher in the SSA-PMG groups than the unselected groups (Fig. 4B and Fig. S5C–S5D). The result was further confirmed by deep sequencing (Fig. 4C–D). Collectively, the above results indicated that the SSA-PMG surrogate reporter facilitated the efficient selection and enrichment of genetically modified cells with target gene knockout.

3.5. Detection of CRISPR/Cas9-mediated knock-in efficiency in cells enriched by the HDR-PMG surrogate reporter

Our previous work [34] and the above results have all demonstrated that the SSA-based surrogate reporter system can be used to enrich genomically modified cells, especially for the cells harboring NHEJ-based indels and target gene knockout. To specifically enrich the HDR-based genetically modified cells, we next developed the HDR-PMG reporter. HEK293T cells were co-transfected with the pGOI-sgRNA/Cas9 expression vectors, the corresponding reporters (pGOI-SSA-PMG or pGOI-HDR-PMG) and donor plasmids (pD-GOI) (Fig. 5A, E–F). Three days after transfection, the cells were sorted by FACS or puromycin-selection for subsequent digestion assays. After flow sorting, the HDR-PMG group yielded 21.4 %, 17.5 % and 11.7 % HDR-based editing efficiencies at the h*CCR5*, h*EMX1*, and h*NUDT5* loci, respectively, which corresponded to 10.2–10.3- and 5.3-fold of the unsorted groups (Fig. 5B, G–H). Consistently, we found that the HDR-based editing efficiency in the HDR-PMG groups ranged from 23.5 % to 36.5 % after puromycin selection and exhibited a significant increase as 10.7- to 17.4-fold of that in the unsorted groups (Fig. 5B, G–H). The impairment of HDR-based editing efficiency in cells enriched by the HDR-PMG reporter was further confirmed by deep sequencing assays (Fig. 5C–D and I–J). Interestingly, although the SSA-PMG surrogate reporter was more



(caption on next page)

Fig. 2. Enrichment of genetically modified cells with the SSA-PMG and HDR-PMG reporters.
(A) Schematic of the enrichment of genetically modified cells using the SSA-PMG or HDR-PMG reporters. (B) Flow cytometry analysis for the detection of the self-repairing efficiency of the SSA-PMG or HDR-PMG reporters. (C) Fluorescence microscopy of transfected cells for enriched $EGFP^{+}$ and $Puro^{R+}$ cells 3 days after the puromycin selection. (D) T7E1 assay result for indel frequencies in unselected and selected cells ($EGFP^{+}$ or $Puro^{R+}$). Arrows indicate the expected positions of DNA bands cleaved by mismatch-sensitive T7E1. Mutation frequencies (indels, %) at the bottom were calculated by measuring the band intensities. (E) Sequencing results of indels in unselected and selected cells (EGFP-selected or puromycin-selected) transfected with the phAAVS1-SSA-PMG reporter and phAAVS1-sgRNA/Cas9. Dashes and lower-case letters in red indicate deleted and inserted base pairs, respectively (the inserted or deleted base pairs are numbered in the parentheses). Mutation frequencies were calculated by dividing the number of mutant clones by the number of total clones. (F) Experimental schematic of the editing at h*AAVS*1 locus. The donors contained a left homology arm (LHA) (1036 bp) and a right homology arm (RHA) (1057 bp), and the PAM site was replaced by an *Xba* I restriction enzyme recognition site. (G) The effect of HDR-PMG surrogate reporter on HDR repair frequencies at the h*AAVS*1 locus in HEK293T cells was detected by restriction enzyme digestion assay. Scale bar = 100 μm.

sensitive to monitor the CRISPR/Cas9 nuclease activity, the HDR-PMG surrogate reporter appeared more suitable for the enrichment of cells harboring gene knock-in edits. Our data demonstrated that high knock-in efficiency could be achieved by applying the HDR-PMG reporter.

3.6. Enhancement of HDR-based knock-in editing by the collaboration of HDR-PMG surrogate reporter and other strategies

To further enhance the HDR-based knock-in editing, the HDR-PMG surrogate reporter was combined with several other strategies. It was previously shown that HDR efficiency could be improved by using optimized plasmid donors containing two CRISPR/Cas targets flanking the recombination template, which is also known as double-cut donors [30]. Hence, we first compared the different types of donor templates on HDR efficiency, including plasmid donor, linear donor and double-cut plasmid donor, at two genomic sites in HEK293T cells (Fig. 6A). After puromycin selection, the linear donors and double-cut donors exhibited higher HDR efficiency than the plasmid donors (Fig. 6B–E, Table S6). In addition, we tested the effect of the HA length of the ssODN templates on HDR efficiency (Fig. 6F). The results indicated that there was a correlation between ssODN length and HDR efficiency, with longer ssODN resulting in higher HDR efficiency, and that the 130-nt ssODN template exhibited higher efficiency than plasmid donors (Fig. 6G–J, Table S7).

Higher HDR repair efficiency can be achieved by using small molecules or proteins that either inhibit NHEJ or prolong the S and G2 phases of cell cycle [29]. In the attempt to further improve HDR efficiency, we constructed a bifunctional vector for the co-expression of CRISPR/Cas9 and HDR enhancing factors (yRad52 or Ad4E1B-E4orf6) [23]. Digestion assays and deep sequencing analysis revealed that co-expression of yRad52 significantly improved HDR frequency at the h*EMX*1 and h*NUDT*5 loci. Contrary to expectations, ectopic expression of Ad4E1B-E4orf6 contributed only minor improvements in HDR efficiency at the h*EMX*1 locus, but resulted in a modest decrease at the h*NUDT*5 locus (Fig. 6K–N, Table S8). The HDR-based editing efficiency with Scr7 exhibited a modest decrease at the h*EMX*1 and h*NUDT*5 locus. As expected, treatment with nocodazole increased the HDR-based editing events at both genomic loci (Fig. 6O–R, Table S9). Overall, these experiments established that the efficiency of HDR-based knock-in editing could be further improved by the collaborative usage of our HDR-PMG surrogate reporter and other HDR enhancing strategies.

3.7. Application of the SSA-PMG and HDR-PMG surrogate reporters for different CRISPR/Cas systems and in primary mammalian cells

To further expand the application of these surrogate reporters, we next tested whether SSA-PMG and HDR-PMG could be adapted to other CRISPR/Cas systems. HEK293T cells were co-transfected with the CRISPR/Cas expression plasmids (gRNA/SlugCas9 or gRNA/AsCas12a) targeting the human *VEGFA* gene together with the corresponding SSA-PMG and HDR-PMG reporters. Three days after transfection, a substantial number of EGFP-positive cells were observed under fluorescence microscopy (Fig. 7A–B), indicating that the two surrogate reporters are also applicable to different CRISPR/Cas systems. This result was also confirmed by flow cytometry assays (Figs. 7C–E).

Primary cells have been gradually becoming the preferred tool in cellular and molecular studies for the regulation mechanisms of animal growth and development. Therefore, we further evaluated our SSA-PMG and HDR-PMG reporters in cells derived from a different species, including mESCs, PEFs, and BMECs. The cells were respectively co-transfected with the CRISPR/Cas9 expression plasmid and the SSA-PMG reporter without a donor or the HDR-PMG reporter with the donor. Three days after transfection, the SSA-PMG and HDR-PMG reporters in the transfected cells were observed to be successfully repaired by restored EGFP fluorescence, indicating that the two reporters can also be utilized to enrich genetically modified primary cells (Fig. S6). We further examined the genome editing efficiency in the primary goat myoblasts enriched by the SSA-PMG and HDR-PMG reporters (Fig. 7F–H). The T7E1 assays revealed that the knockout efficiency of SSA-PMG group was 26.6 %, significantly higher than the 14.7 % in the pPuro-T2A-EGFP group after puromycin selection for 3 days (Fig. 7G). The subsequent sequencing analysis further corroborated this result (Fig. 7H). In addition, the digestion assays revealed four knock-in events out of 45 clones (4/45) from the HDR-PMG group, but the knock-in events were undetectable in the pPuro-T2A-EGFP screening group (Fig. 7J–7K). Taken together, our surrogate reporters were effective for the enrichment of genetically modified primary mammalian cells.

4. Discussion

In recent years, significant progress has been made in the field of gene editing. Emerging gene editing technologies have been successfully applied for the genetic modification in various cell types and organisms. In this study, we developed two novel screening traffic light reporters, SSA-PMG and HDR-PMG, which permit simple and efficient enrichment of knockout and knock-in genetically modified cells, respectively, through FACS or puromycin selection. A key aspect of the surrogate reporters is the reporter construct with dual-reporter genes and the fluorescent marker gene assembled in the same open reading frame driven by a single CMV promoter, which contributed to reduce the plasmid size, and make it much easier for cell transfection. The SSA-PMG and HDR-PMG reporters were proven effective for the enrichment of genetically modified cells at different endogenous loci and in different cell lines. In addition, we also validated the applicability of the reporters in several primary mammalian cells. Thus, the SSA-PMG and HDR-PMG reporters developed in this study expand the current surrogate reporter toolbox for genome editing.

Several similar surrogate reporters, based on the different DSB repair pathways, have been developed for *in vitro* analysis of nuclease activity as well as the enrichment of genetically modified cells (Table 2). Compared with other reporters, an obvious advantage of our PMG surrogate reporters are the puromycin-resistance and EGFP dual-reporter genes and the *mCherry* fluorescent marker gene assembled in the same expression cassette. The relatively small size not only provides the advantage in gene delivery but also makes it structurally and functionally more flexible and refined. The resistance and fluorescence dual-reporter design enables both monitoring of nuclease activity and screening of genome-edited cells [21,34]. Since the PMG reporters are based on the correction of the truncated puromycin-resistance cassette,

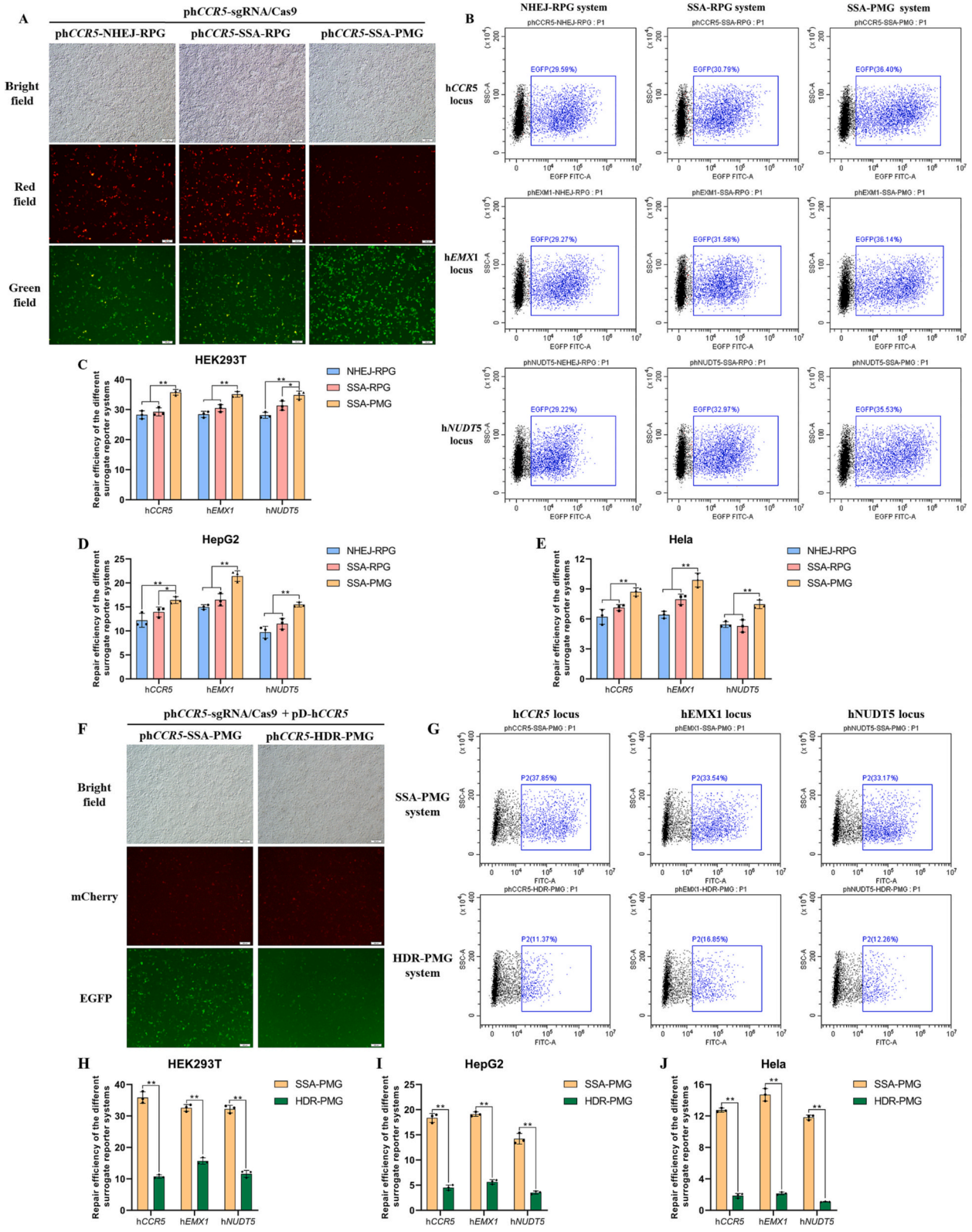


Fig. 3. Comparison of the self-repairing efficiency and sensitivity of different surrogate reporters.
(A and B) Fluorescence microscopy (A) and flow cytometry analysis (B) were used to compare the self-repairing efficiency and sensitivity of different surrogate reporters. (C and E) Self-repairing efficiency of different surrogate reporters in HEK293T (C), HepG2 (D) and Hela cells (E) by flow cytometry analysis. (F and G) The repair efficiency and sensitivity efficiency of SSA-PMG and HDR-PMG vector systems were confirmed by fluorescence microscopy (F) and flow cytometry analysis (G). (H-J) Comparison of self-repairing efficiency and sensitivity of SSA-PMG and HDR-PMG reporters in HEK293T (H), HepG2 (I) and Hela cells (J). Data are shown by mean ± SEM of three independent experiments. * $P < 0.05$; ** $P < 0.01$. (Scale bar = 100 μm).

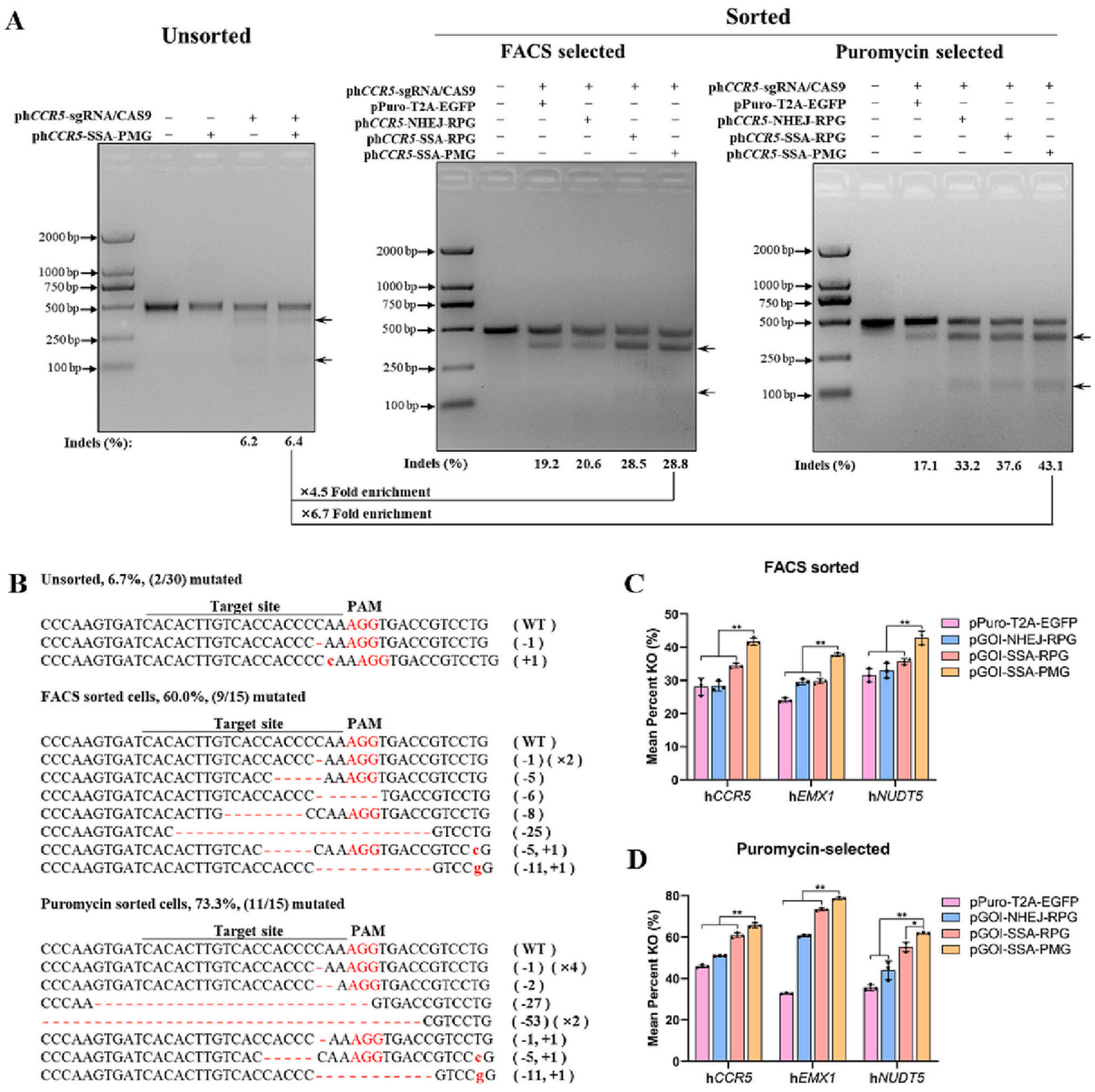


Fig. 4. CRISPR/Cas9-mediated knockout efficiency in cells enriched by the SSA-PMG surrogate reporter.
(A) T7E1 assay results for the indels in selected cells ($EGFP^+$ or $Puro^{R+}$) using different surrogate reporters. Arrows indicate the expected positions of DNA bands cleaved by mismatch-sensitive T7E1. Mutation frequencies (Indels, %) at the bottom were calculated by measuring the band intensities. (B) Sequencing results of indels in SSA-PMG based unselected, FACS-selected and puromycin-selected cells. *Dashes* and *lower-case letters* in red indicate deleted and inserted base pairs, respectively (the inserted or deleted base pairs are numbered in the parentheses). Mutation frequencies were calculated by dividing the number of mutant clones by the number of total clones. (C and D) Deep sequencing results of the knockout efficiency in cells screened by flow cytometry (C) or puromycin (D) using different surrogate reporters. Data are shown by mean ± SEM of three independent experiments. * $P < 0.05$; ** $P < 0.01$.

another advantage is that the nuclease-positive cells could also be easily isolated by puromycin selection. Notably, the design is simple, flexible and can be modified by altering markers to provide a choice of preferred screening modalities. The fluorescence-based screening strategy is usually limited by the flow cytometer equipment capabilities, contamination and poor cell growth, while the antibiotic resistance-based screening strategy offers an additional and perhaps more practical route.

Efficient DSB repair is crucial for the maintenance of genome integrity and cell survival. Although NHEJ has been widely regarded as a predominant DSB repair pathway in mammalian cells, SSA represents a major homologous recombination pathway for repairing DSBs flanked by direct repeats [36,37]. We have shown previously that repair efficiency of the SSA-based reporter increases with the length of the direct repeats to a certain extent. The SSA-based reporter with a direct repeat length of 200 bp is more sensitive and efficient than the NHEJ-based reporter [34]. Therefore, the SSA-PMG reporter was constructed with the direct repeat length of 200 bp.

CRISPR/Cas-mediated HDR-based gene editing has great significance for gene function studies and molecular therapy. However, HDR occurs infrequently (typically ~0.1–5 %) in mammalian cells, impeding its application in precise editing and gene knock-in [38]. The strategy, termed co-targeting with selection, increased the probability of isolation

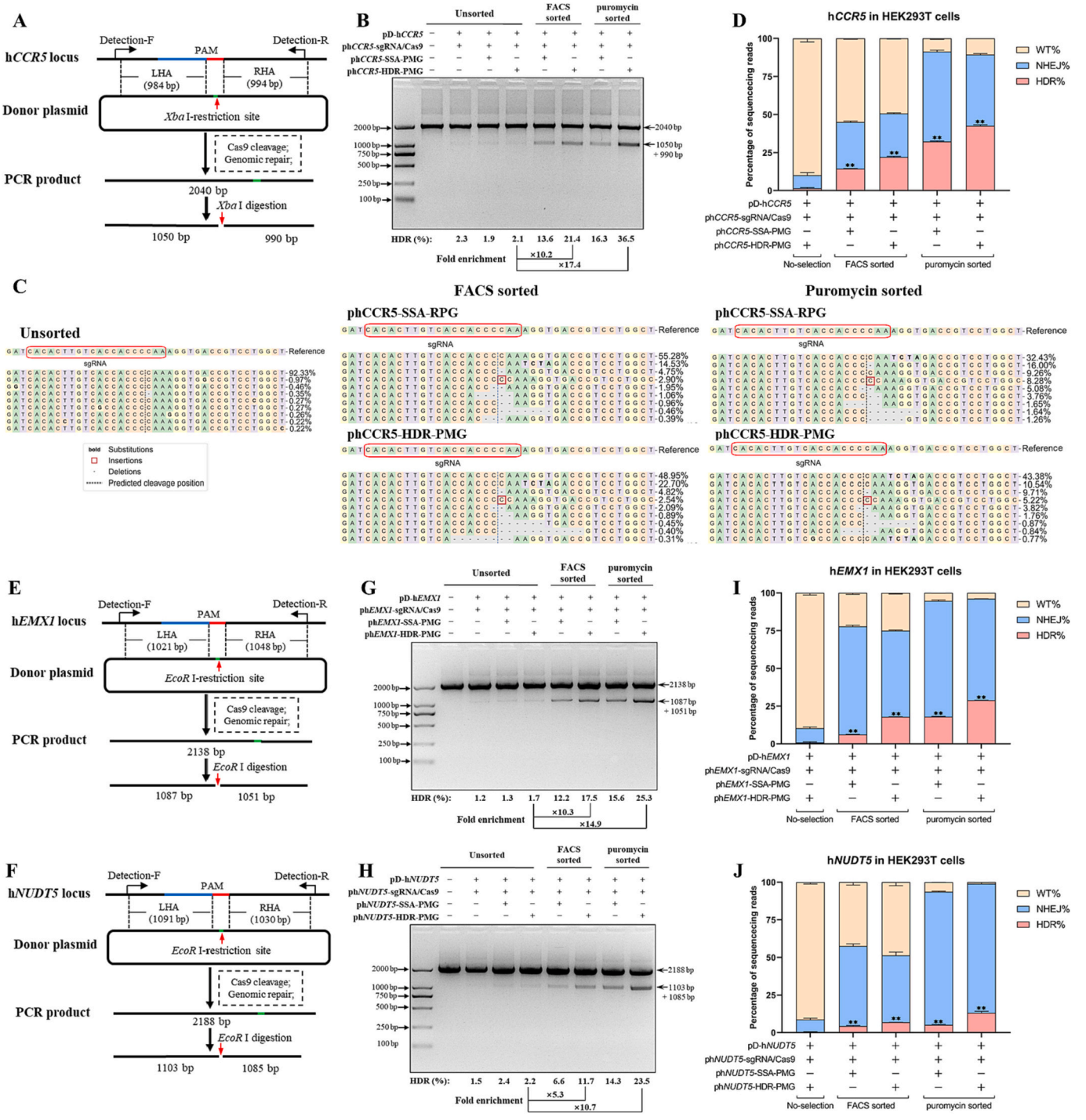


Fig. 5. CRISPR/Cas9-mediated knock-in efficiency in cells enriched by the HDR-PMG surrogate reporter.
(A) Schematic of the knock-in editing at h*CCR*5 gene locus. The donors contained a left homology arm (LHA) and a right homology arm (RHA), and the PAM site was replaced by an *Xba* I restriction enzyme recognition site. (B) Comparison of the HDR-based knock-in efficiency at the h*CCR*5 locus in cells enriched by the SSA-PMG and HDR-PMG reporters. (C) CRISPResso2 analysis of h*CCR*5 gene editing. (D) Deep sequencing result of the knock-in efficiency at the h*CCR*5 loci in cells screened by flow cytometry (C) or puromycin (D) using the SSA-PMG and HDR-PMG reporters. (E and F) Schematic of the knock-in editing at the h*EMX*1 (E) and h*NUDT*5 (F) gene loci. The PAM site was replaced by an *EcoR* I restriction enzyme recognition site. (G and H) Comparison of the knock-in editing efficiency at the h*EMX*1 (G) and h*NUDT*5 (H) gene loci. (I and J) Deep sequencing result of the knock-in efficiency at the h*EMX*1(I) and h*NUDT*5 (J) loci. Data are shown by mean ± SEM of three independent experiments. * $P < 0.05$; ** $P < 0.01$.

and enrichment of HDR-proficient cells with an HDR-specific surrogate reporter. For efficient HDR-based editing, the length of the homologous arms and the HDR donor types are critical considerations. We previously showed that a donor with HA length of >800 bp (800-bp left and 800-bp right) exhibited relatively higher HDR efficiency [23]. Consistent with the previous research, we found that linear and double-cut donors mediated higher HDR efficiencies than plasmid donors. Additionally, we observed that the 130-nucleotide ssODN template had the highest HDR efficiency among all types of donors. These data suggested that the enrichment of knock-in cells with our HDR-PMG reporter can be further

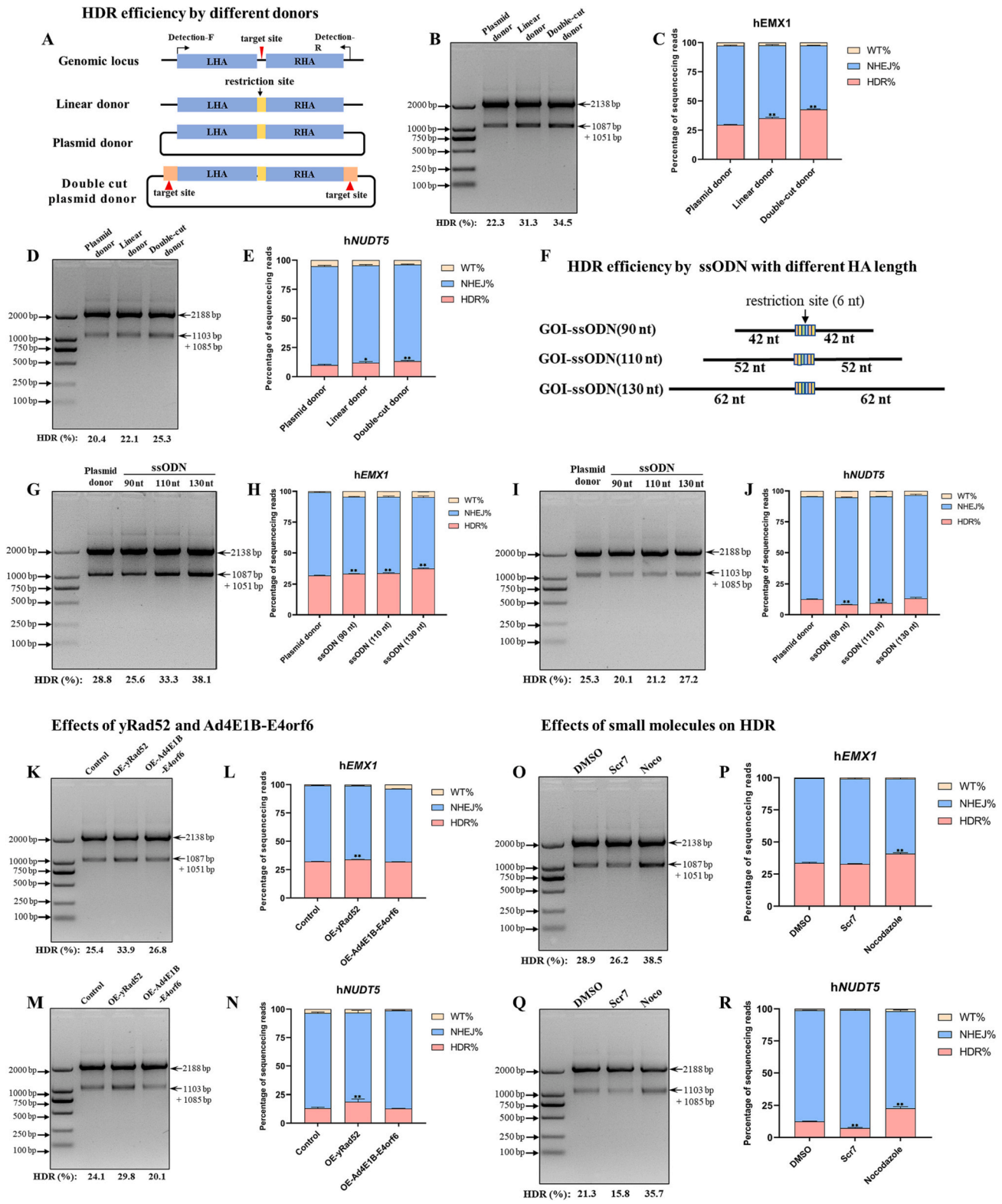


Fig. 6. Improved HDR-based knock-in efficiency by the collaboration of HDR-PMG surrogate reporter and other strategies
(A) Schematic of the different donor modalities used in this study. (B-E) The effects of the plasmid donor, linear donor and double-cut linear donor on HDR-based editing efficiency at the h*EMX*1 (B and C) and h*NUDT*5 (D and E) loci in HEK293T cells, detected by restriction enzyme digestion (B and D) and deep sequencing (C and E). (F) Schematic of ssODN donors with different length of homologous arms used in this study. (G-J) The effects of different ssODN donors (with 90-nt, 110-nt, and 130-nt lengths) on HDR-based editing efficiency at the h*EMX*1 (G and H) and h*NUDT*5 (I and J) loci, detected by restriction enzyme digestion (G and I) and deep sequencing (H and J). (K—N) The effects of yRad52 or Ad4E1B-E4orf6 on HDR-based editing efficiency at the h*EMX*1 (K and L) and h*NUDT*5 (M and N) loci, detected by restriction enzyme digestion (K and M) and deep sequencing (L and N). (O-R) The effects of small molecules Scr7 and nocodazole on HDR-based editing efficiency at the h*EMX*1 (O and P) and h*NUDT*5 (Q and R) loci, detected by restriction enzyme digestion (O and Q) and deep sequencing (P and R). Data are shown by mean ± SEM of three independent experiments. * $P < 0.05$; ** $P < 0.01$.

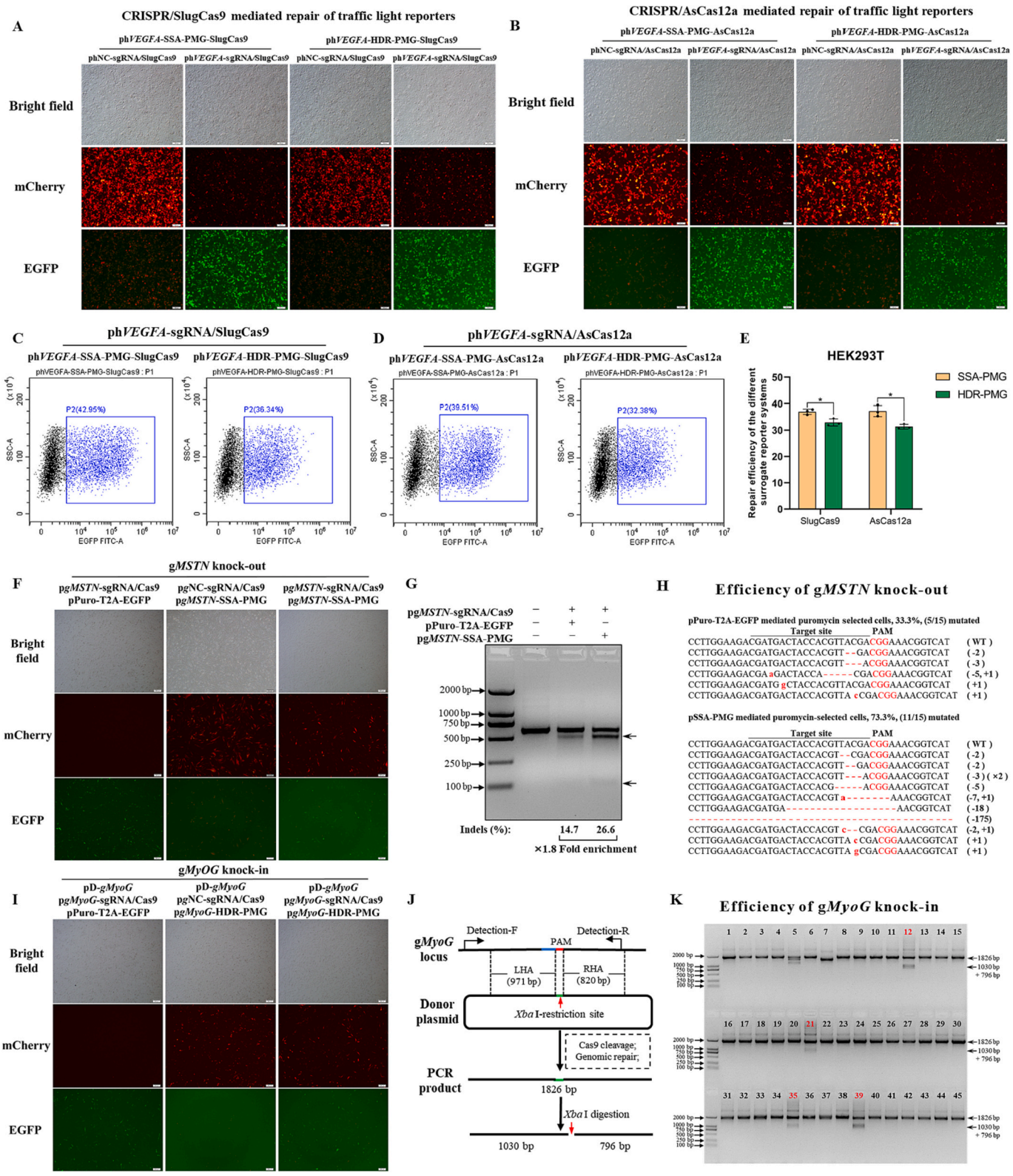


Fig. 7. Application of the SSA-PMG and HDR-PMG surrogate reporters for different CRISPR/Cas systems and in primary mammalian cells.
(A-E) Fluorescence microscopy (A and B) and flow cytometry analysis (C-E) were used to compare the repair efficiency and sensitivity of SSA-PMG and HDR-PMG reporters mediated by CRISPR/SlugCas9 and CRISPR/AsCas12a systems. (Scale bar = 100 μm). (F) Fluorescence microscopy examination was used to analyze the repair efficiency of the SSA-PMG reporter in primary goat myoblasts. (Scale bar = 200 μm). (G-H) T7E1 cleavage assays(G) and DNA sequencing (H) for the mutation frequencies at the g*MSTN* gene locus in puromycin-selected cells. (I) The repair efficiency of HDR-PMG reporter in primary goat myoblasts was observed under fluorescence microscopy. (Scale bar = 200 μm). (J) Schematic of HDR-based editing at the g*MyoG* locus. The donors contained LHA (971 bp) and RHA (820 bp), and the PAM site was replaced by an *Xba* I restriction enzyme recognition site. (K) The result of the restriction enzyme digestion assay for the HDR-based editing efficiency at the g*MyoG* locus in the HDR-PMG selected primary goat myoblasts. Data are shown by mean ± SEM of three independent experiments. * $P < 0.05$; ** $P < 0.01$.

Table 2
Different traffic light surrogate reporters developed for gene editing.

Vector name	Molecular design	Marker gene	Reporter gene	M-R expression manner	Reporter light on by	Marker light off?	Selecting methods		Applicable editing types	Validated in cell types	Ref
							FACS	Drug screening			
Surrogate reporter system		mRFP	EGFP	Tandem	NHEJ	No	Yes	No	Knockout	HEK293T	[42]
Surrogate reporter system		iRFP	EGFP	Tandem	NHEJ	No	Yes	No	Knockout	HEK293T, NALM-6	[43]
Surrogate reporter system		mRFP	EGFP, Hygromycin	Tandem	NHEJ	No	Yes	Yes	Knockout	HEK293T, Huh 7.5 cells	[44]
pBSR		mCherry	Blasticidin S	Tandem	NHEJ	No	No	Yes	Knockout	CH12F3–2	[45]
pRep.eGFP		DsRed	EGFP	Independent	SSA	No	Yes	No	Knockout	HEK293T	[46]
C-check		AsRed	EGFP	Independent	SSA	No	Yes	No	Knockout	HEK293T, MCF7, NHDF, PPF	[25]
HDR-based surrogate reporter		DsRed	EGFP	Independent	HDR	No	Yes	No	Knock-in	HEK293T, PK15	[27]
NHEJ-RPG		DsRed	Puromycin, EGFP	Independent	NHEJ	No	Yes	Yes	Knockout	HEK293T	[34]
SSA-RPG		DsRed	Puromycin, EGFP	Independent	SSA	No	Yes	Yes	Knockout	HEK293T	[34]
SSA-PMG		mCherry	Puromycin, EGFP	Tandem	SSA	Yes	Yes	Yes	Knockout	HEK293T, HepG2, Hela, mESCs, pEFs, BEMCs, Primary goat myoblasts	This study
HDR-PMG		mCherry	Puromycin, EGFP	Tandem	HDR	Yes	Yes	Yes	Knock-in	HEK293T, HepG2, Hela, Primary goat myoblasts	This study

improved.

Recent studies have demonstrated that the overexpression of yRad52 and Ad4E1B-E4orf6 could improve the HDR-based gene knock-in efficiency [27,39]. In addition, small molecules such as Scr7 and nocodazole have also been reported to enhance the CRISPR/Cas-mediated HDR efficiency [23]. After experimentally examining these factors, we found that overexpression of yRad52 or added nocodazole increased HDR events. Contrary to expectations, neither Ad4E1B-E4orf6 nor Scr7 had a noticeable positive impact on the HDR efficiency. One possible explanation for the discrepancy between our results and previous studies is that our surrogate reporter strategy had already achieved high-level HDR events in the selected cells, which masks the subtle changes caused by the additional measures. However, further studies are still required to understand the detailed mechanisms.

Novel precise and efficient programmable nucleases for genetic engineering are continuously emerging nowadays. However, different nucleases exhibited various efficiencies in the process of gene editing, most remains to be optimized for further improvement. Fortunately, surrogate reporters with compatibility provide efficient tools for the activity validation of different nucleases. In the current study, the SSA-PMG and HDR-PMG reporters were proved applicable to different CRISPR/Cas systems. Moreover, most of the research using CRISPR/Cas9 for genome editing has mainly focused on immortalized cell lines, while primary cells and embryonic stem cells were rarely investigated [40,41]. To date, only a handful of studies have applied CRISPR/Cas9 to edit primary cells isolated from animal tissues. The gene editing in primary mammalian cells is usually restricted by their poor transfection efficiency and limited lifespan, which is also dependent on the positive selection and enrichment strategies. Hence, we further tested our SSA-PMG and HDR-PMG reporters in several primary mammalian cells, including mESCs, PEFs and BMECs, and successfully generated clonogenic primary goat myoblasts with targeted gene knockout and knock-in. We believe that the surrogate reporter-based screening strategy has great potential to facilitate the gene editing of primary cells.

5. Conclusion

In summary, we developed the novel SSA-PMG and HDR-PMG surrogate reporters, which provide robust and powerful approaches for enriching CRISPR/Cas-mediated knockout and knock-in cells, respectively. We envision that the two reporters will facilitate the use of programmable nucleases in both basic research and biotechnology applications.

CRediT authorship contribution statement

Ming Lyu: Investigation, Methodology; Data curation; Data analysis; Visualization, Writing - original draft, Writing - review & editing. Yongsen Sun, Nana Yan, Qiang Chen, Xin Wang and Zehui Wei: Methodology; Data curation; Data analysis; Zhiying Zhang, and Kun Xu: Methodology; Funding acquisition; Project administration; Writing - revision & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank all the members of the Animal Genome Editing Laboratory of the College of Animal Science and Technology, Northwest A&F University who contributed their efforts to these experiments. This study was supported by grants from the National Natural Science Foundation of China (31702099).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2023.124926.

References

[1] F.D. Urnov, E.J. Rebar, M.C. Holmes, H.S. Zhang, P.D. Gregory, Genome editing with engineered zinc finger nucleases., nature reviews, Genetics 11 (9) (2010) 636–646.
[2] T.A. Mcmurrough, C.M. Brown, K. Zhang, G. Hausner, M.S. Junop, G.B. Gloor, D. R. Edgell, Active site residue identity regulates cleavage preference of laglidadg homing endonucleases, Nucleic Acids Res. 46 (22) (2018) 11990–12007.
[3] J.K. Joung, J.D. Sander, Talens: a widely applicable technology for targeted genome editing, Nat. Rev. Mol. Cell Biol. 14 (1) (2013) 49–55.
[4] B. Wiedenheft, S.H. Sternberg, J.A. Doudna, Rna-guided genetic silencing systems in bacteria and archaea, Nature 482 (7385) (2012) 331–338.
[5] J. Tan, F. Zhang, D. Karcher, R. Bock, Engineering of high-precision base editors for site-specific single nucleotide replacement, Nat. Commun. 10 (1) (2019) 439.
[6] S.A. Hosseini, A. Salehifard Jouneghani, M. Ghatrehsamani, H. Yaghoobi, F. Elahian, S.A. Mirzaei, Crispr/cas9 as precision and high-throughput genetic engineering tools in gastrointestinal cancer research and therapy, Int. J. Biol. Macromol. 223 (Pt A) (2022) 732–754.
[7] D.O. Reilly, Z.J. Kartje, E.A. Ageely, E. Malek-Adamian, M. Habibian, A. Schofield, C.L. Barkau, K.J. Rohilla, L.B. Derossett, A.T. Weigle, M.J. Damha, K.T. Gagnon, Extensive crispr rna modification reveals chemical compatibility and structure-activity relationships for cas9 biochemical activity, Nucleic Acids Res. 47 (2) (2018) 546–558.
[8] J.E. Garneau, M. Dupuis, M. Villion, D.A. Romero, R. Barrangou, P. Boyaval, C. Fremaux, P. Horvath, A.H. Magadán, S. Moineau, The crispr/cas bacterial immune system cleaves bacteriophage and plasmid dna, Nature 468 (7320) (2010) 67–71.
[9] G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9-crrna ribonucleoprotein complex mediates specific dna cleavage for adaptive immunity in bacteria, Proc. Natl. Acad. Sci. U. S. A. 109 (39) (2012) E2579–E2586.
[10] B. Zetsche, J.S. Gootenberg, O.O. Abudayyeh, I.M. Slaymaker, K.S. Makarova, P. Essletzbichler, S.E. Volz, J. Joung, J. van der Oost, A. Regev, E.V. Koonin, F. Zhang, Cpf1 is a single rna-guided endonuclease of a class 2 crispr-cas system, Cell 163 (3) (2015) 759–771.
[11] D. Baisya, A. Ramesh, C. Schwartz, S. Lonardi, I. Wheeldon, Genome-wide functional screens enable the prediction of high activity crispr-cas9 and -cas12a guides in yarrowia lipolytica, Nat. Commun. 13 (1) (2022) 922.
[12] L. Cong, F.A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P.D. Hsu, X. Wu, W. Jiang, L.A. Marraffini, F. Zhang, Multiplex genome engineering using crispr/cas systems, Science vol. 339 (6121) (2013) 819–823. New York, N.Y.
[13] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, E. Charpentier, A programmable dual-rna-guided dna endonuclease in adaptive bacterial immunity, Science vol. 337 (6096) (2012) 816–821. New York, N.Y.
[14] A. Mehta, J.E. Haber, Sources of dna double-strand breaks and models of recombinational dna repair, Cold Spring Harb. Perspect. Biol. 6 (9) (2014), a16428.
[15] M.R. Lieber, The mechanism of double-strand dna break repair by the nonhomologous dna end-joining pathway, Annu. Rev. Biochem. 79 (1) (2010) 181–211.
[16] F. Aymard, B. Bugler, C.K. Schmidt, E. Guillou, P. Caron, S. Briois, J.S. Iacovoni, V. Daburon, K.M. Miller, S.P. Jackson, G. Legube, Transcriptionally active chromatin recruits homologous recombination at dna double-strand breaks, Nat. Struct. Mol. Biol. 21 (4) (2014) 366–374.
[17] M.C. Canver, M. Haeussler, D.E. Bauer, S.H. Orkin, N.E. Sanjana, O. Shalem, G. C. Yuan, F. Zhang, J.P. Concordet, L. Pinello, Integrated design, execution, and analysis of arrayed and pooled crispr genome-editing experiments, Nat. Protoc. 13 (5) (2018) 946–986.
[18] A. Sallmyr, A.E. Tomkinson, Repair of dna double-strand breaks by mammalian alternative end-joining pathways, J. Biol. Chem. 293 (27) (2018) 10536–10546.
[19] E.L. Ivanov, N. Sugawara, J. Fishman-Lobell, J.E. Haber, Genetic requirements for the single-strand annealing pathway of double-strand break repair in saccharomyces cerevisiae, Genetics (Austin) 142 (3) (1996) 693–704.
[20] R. Ceccaldi, B. Rondinelli, A.D.D. Andrea, Repair pathway choices and consequences at the double-strand break, Trends Cell Biol. 26 (1) (2016) 52–64.
[21] R. Kuhar, K.S. Gwiazda, O. Humbert, T. Mandt, J. Pangallo, M. Brault, I. Khan, N. Maizels, D.J. Rawlings, A.M. Scharenberg, M.T. Certo, Novel fluorescent genome editing reporters for monitoring dna repair pathway utilization at endonuclease-induced breaks, Nucleic Acids Res. 42 (1) (2014), e4.
[22] H. Kim, E. Um, S. Cho, C. Jung, H. Kim, J. Kim, Surrogate reporters for enrichment of cells with nuclease-induced mutations, Nat. Methods 8 (11) (2011) 941–943.
[23] N. Yan, Y. Sun, Y. Fang, J. Deng, L. Mu, K. Xu, J.S. Mymryk, Z. Zhang, A universal surrogate reporter for efficient enrichment of crispr/cas9-mediated homology-directed repair in mammalian cells, Molecular therapy. Nucleic acids 19 (2020) 775–789.
[24] S.W. Cho, S. Kim, J.M. Kim, J. Kim, Targeted genome engineering in human cells with the cas9 rna-guided endonuclease, Nat. Biotechnol. 31 (3) (2013) 230–232.
[25] Y. Zhou, Y. Liu, D. Hussmann, P. Brøgger, R.A. Al-Saaidi, S. Tan, L. Lin, T. S. Petersen, G.Q. Zhou, P. Bross, L. Aagaard, T. Klein, S.G. Rønn, H.D. Pedersen, L. Bolund, A.L. Nielsen, C.B. Sørensen, Y. Luo, Enhanced genome editing in mammalian cells with a modified dual-fluorescent surrogate system, Cellular and molecular life sciences: CMLS 73 (13) (2016) 2543–2563.
[26] S.B. Jung, C.Y. Lee, K. Lee, K. Heo, S.H. Choi, A cleavage-based surrogate reporter for the evaluation of crispr–cas9 cleavage efficiency, Nucleic Acids Res. 49 (15) (2021), e85.
[27] S. Shao, C. Ren, Z. Liu, Y. Bai, Z. Chen, Z. Wei, X. Wang, Z. Zhang, K. Xu, Enhancing crispr/cas9-mediated homology-directed repair in mammalian cells by expressing saccharomyces cerevisiae rad52, Int. J. Biochem. Cell Biol. 92 (2017) 43–52.
[28] L. Wang, L. Yang, Y. Guo, W. Du, Y. Yin, T. Zhang, H. Lu, Enhancing targeted genomic dna editing in chicken cells using the crispr/cas9 system, PLoS One 12 (1) (2017), e169768.
[29] S. Lin, B.T. Staahl, R.K. Alla, J.A. Doudna, Enhanced homology-directed human genome engineering by controlled timing of crispr/cas9 delivery, Elife 3 (2014), e4766.
[30] J. Zhang, X. Li, G. Li, W. Chen, C. Arakaki, G.D. Botimer, D. Baylink, L. Zhang, W. Wen, Y. Fu, J. Xu, N. Chun, W. Yuan, T. Cheng, X. Zhang, Efficient precise knockin with a double cut hdr donor after crispr/cas9-mediated double-stranded dna cleavage, Genome Biol. 18 (1) (2017) 35.
[31] Z. Hu, C. Zhang, S. Wang, S. Gao, J. Wei, M. Li, L. Hou, H. Mao, Y. Wei, T. Qi, H. Liu, D. Liu, F. Lan, D. Lu, H. Wang, J. Li, Y. Wang, Discovery and engineering of small slugcas9 with broad targeting range and high specificity and activity, Nucleic Acids Res. 49 (7) (2021) 4008–4019.
[32] M.A. Moreno-Mateos, J.P. Fernandez, R. Rouet, C.E. Vejnar, M.A. Lane, E. Mis, M. K. Khokha, J.A. Doudna, A.J. Giraldez, Crispr-cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing, Nat. Commun. 8 (1) (2017) 2024.
[33] X. Yao, X. Wang, J. Liu, X. Hu, L. Shi, X. Shen, W. Ying, X. Sun, X. Wang, P. Huang, H. Yang, Crispr/cas9 - mediated precise targeted integration in vivo using a double cut donor with short homology arms, Ebiomedicine 20 (2017) 19–26.
[34] C. Ren, K. Xu, Z. Liu, J. Shen, F. Han, Z. Chen, Z. Zhang, Dual-reporter surrogate systems for efficient enrichment of genetically modified cells. 72 (2015) 2763–2772. https://https://doi.org/10.1007/s00018-015-1874-6.
[35] M. Lyu, X. Wang, X. Meng, H. Qian, Q. Li, B. Ma, Z. Zhang, K. Xu, Chi-mir-487b-3p inhibits goat myoblast proliferation and differentiation by targeting irs1 through the irs1/pi3k/akt signaling pathway, Int. J. Mol. Sci. 23 (1) (2022) 115.
[36] S.W. Morrical, Dna-pairing and annealing processes in homologous recombination and homology-directed repair, Cold Spring Harb. Perspect. Biol. 7 (2) (2015), a16444.
[37] L.S. Symington, Role of rad52 epistasis group genes in homologous recombination and double-strand break repair, Microbiology and molecular biology reviews: MMBR 66 (4) (2002) 630–670.
[38] J. Xie, W. Ge, N. Li, Q. Liu, F. Chen, X. Yang, X. Huang, Z. Ouyang, Q. Zhang, Y. Zhao, Z. Liu, S. Gou, H. Wu, C. Lai, N. Fan, Q. Jin, H. Shi, Y. Liang, T. Lan, L. Quan, X. Li, K. Wang, L. Lai, Efficient base editing for multiple genes and loci in pigs using base editors, Nat. Commun. 10 (1) (2019) 2852.
[39] V.T. Chu, T. Weber, B. Wefers, W. Wurst, S. Sander, K. Rajewsky, R. Kühn, Increasing the efficiency of homology-directed repair for crispr-cas9-induced precise gene editing in mammalian cells, Nat. Biotechnol. 33 (5) (2015) 543–548.
[40] P.D. Hsu, E.S. Lander, F. Zhang, Development and applications of crispr-cas9 for genome engineering, Cell 157 (6) (2014) 1262–1278.
[41] D.B.T. Cox, R.J. Platt, F. Zhang, Therapeutic genome editing: prospects and challenges, Nat. Med. 21 (2) (2015) 121–131.
[42] S. Ramakrishna, S.W. Cho, S. Kim, M. Song, R. Gopalappa, J. Kim, H. Kim, Surrogate reporter-based enrichment of cells containing rna-guided cas9 nuclease-induced mutations, Nat. Commun. 5 (1) (2014) 3378.
[43] W. Liu, K. Völse, D. Senft, I. Jeremias, A reporter system for enriching crispr/cas9 knockout cells in technically challenging settings like patient models, Sci. Rep. 11 (1) (2021) 12649.
[44] H. Kim, M. Kim, G. Wee, C. Lee, H. Kim, J. Kim, Magnetic separation and antibiotics selection enable enrichment of cells with zfn/talen-induced mutations, PLoS One 8 (2) (2013), e56476.
[45] F. Niccheri, R. Pecori, S.G. Conticello, An efficient method to enrich for knock-out and knock-in cellular clones using the crispr/cas9 system, Cell. Mol. Life Sci. 74 (18) (2017) 3413–3423.
[46] K. Xu, C. Ren, Z. Liu, T. Zhang, T. Zhang, D. Li, L. Wang, Q. Yan, L. Guo, J. Shen, Z. Zhang, Efficient genome engineering in eukaryotes using cas9 from streptococcus thermophilus, Cell. Mol. Life Sci. 72 (2) (2015) 383–399.