Molecular Therapy
Nucleic Acids
Original Article

A Universal Surrogate Reporter for Efficient Enrichment of CRISPR/Cas9-Mediated Homology-Directed Repair in Mammalian Cells

Nana Yan,[1,3] Yongsen Sun,[1,3] Yuanyuan Fang,[1] Jingrong Deng,[1] Lu Mu,[1] Kun Xu,[1] Joe S. Mymryk,[2] and Zhiying Zhang[1]

[1]Key Laboratory of Animal Genetics, Breeding and Reproduction of Shaanxi Province, College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi 712100, China; [2]Department of Microbiology & Immunology, Oncology and Otolaryngology, The University of Western Ontario, London, ON N6A 3K7, Canada

CRISPR/Cas9-mediated homology-directed repair (HDR) can be leveraged to precisely engineer mammalian genomes. However, the inherently low efficiency of HDR often hampers to identify the desired modified cells. Here, we developed a novel universal surrogate reporter system that efficiently enriches for genetically modified cells arising from CRISPR/Cas9-induced HDR events (namely, the "HDR-USR" system). This episomally based reporter can be self-cleaved and self-repaired via HDR to create a functional puromycin selection cassette without compromising genome integrity. Co-transfection of the HDR-USR system into host cells and transient puromycin selection efficiently achieves enrichment of HDR-modified cells. We tested the system for precision point mutation at 16 loci in different human cell lines and one locus in two rodent cell lines. This system exhibited dramatic improvements in HDR efficiency at a single locus (up to 20.7-fold) and two loci at once (42% editing efficiency compared to zero in the control), as well as greatly improved knockin efficiency (8.9-fold) and biallelic deletion (35.9-fold) at test loci. Further increases were achieved by co-expression of yeast Rad52 and linear single-/double-stranded DNA donors. Taken together, our HDR-USR system provides a simple, robust and efficient surrogate reporter for the enrichment of CRISPR/Cas9-induced HDR-based precision genome editing across various targeting loci in different cell lines.

INTRODUCTION

Precision genomic editing allows scientists to directly alter gene sequences to investigate gene function and to develop safe and highly precise gene therapy approaches to treat human disease.[1,2] The CRISPR/CRISPR-associated protein-9 nuclease (CRISPR/Cas9) system, which utilizes the Cas9 nuclease in conjunction with a single guide RNA (sgRNA), is empowering researchers to perform precision genome engineering that allows the efficient production of a site-specific DNA double-strand break (DSB) within a genome.[3–5] The DSB can be repaired by two distinct mechanisms, nonhomologous end joining (NHEJ) and homology-directed repair (HDR).[6,7] The NHEJ repair functions in an error-prone pathway characterized by the ligation of DNA ends without end processing, resulting in small frameshift insertion or deletion (indel) mutations at the site of the DSB. In contrast, HDR utilizes an exogenous DNA donor template homologous to the region surrounding the DSB to precisely repair the damage. Importantly, the HDR pathway can be exploited to create base substitutions, insertions of the desired sequence, and precise deletions.[2,7] However, repair by the NHEJ pathway often predominates, restricting the efficiency of the more desirable HDR repair,[4,8] and this negatively impacts the efficiency of precision genome engineering.[9]

To improve the efficiency of precision genome editing, various attempts have been made to enhance HDR. Direct strategies include overexpressing key homologous recombination proteins like Rad51 and Rad52 or the use of HDR agonists.[10,11] The alternative strategy is to inhibit key proteins associated with NHEJ, such as KU or LIG4 by RNAi or using appropriate small-molecule inhibitors.[12–14] Another approach for enhancing HDR is altering cell-cycle parameters to promote and prolong the S and G2 phases, when HDR is most active,[15] or post-translationally restricting the expression of Cas9 during the G1 phase.[16] In addition, the structures of the DNA donor template, including long homology arm (HA) (1 kb or greater) targeting constructs,[17] linear double-stranded donors,[13,18,19] single-stranded oligodeoxynucleotides (ssODNs),[20–22] and double cut donors[23,24] have been used to increase HDR efficiency.

In addition to low HDR efficiency, isolating the rare genetically modified cells from the treated cell population is also technically challenging and time consuming. The initial enrichment method is typically based on the selection of transfection-positive cells with plasmid markers, such as fluorescent proteins or antibiotic-resistance genes.[25–27] These methods select for the plasmid-transfected cell

Received 8 September 2019; accepted 18 December 2019;
https://doi.org/10.1016/j.omtn.2019.12.021.

[3]These authors contributed equally to this work.

Correspondence: Zhiying Zhang, Key Laboratory of Animal Genetics, Breeding and Reproduction of Shaanxi Province, College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi 712100, China.
E-mail: zhangzhy@nwsuaf.edu.cn

clones only and do not ensure the cleavage activity of the artificial nuclease. The most popular surrogate reporters developed for selecting cells bearing the nuclease cleavage events include two types: NHEJ-based surrogate reporters and single-strand annealing (SSA)-based surrogate reporters.[28–30] Both of these surrogate reporters are best suited to enrich for knockout cells, although they can also be used for enrichment of HDR-repaired cells.[11,31] The most common used method for enriching for cells that have undergone precision genome editing is to first to select a knockin cassette that delivers selection marker genes to the target site. This is subsequently followed by a second step that results in the excision of a marker from the knockin cassette during precision editing of the target locus, allowing enrichment by negative selection.[32,33]

Importantly, a recent report demonstrated that HDR-dependent insertion of a selectable marker at one locus coincides with one or more independent HDR-mediated edits at other loci. The concept forms the basis for a strategy has been termed "co-targeting with selection," which can be used to efficiently enrich for HDR-proficient cells.[34,35] However, this approach can result in permanent integration of the selection cassette into the genome of the edited cells. A further refinement to the co-selection strategy is to use a small molecule, ouabain, for marker-free co-selection of NHEJ- or HDR-based editing events by co-targeting the *ATP1A1* locus and the loci of interest.[36] Nevertheless, this approach also requires generation of DSBs at an additional genomic locus besides the primary target.

In this study, we developed and optimized a universal surrogate reporter system specific for efficient enrichment for successful CRISPR/Cas9-mediated HDR without the generation of DSBs at undesired genomic loci, which we have dubbed the "HDR-USR" system. We successfully applied this system in point mutations (including simultaneous dual-locus editing) and fragment indels in mammalian cells. We further examined the effect of different forms of donors, small molecules, and other factors on HDR-USR enrichment efficiency.

RESULTS

A USR System for Enrichment of HDR-Mediated Precision Editing

Based on the phenomenon of "co-targeting with selection" in genome-editing experiments,[34,35] we developed the HDR-USR system. As distinguished from previous approaches using a chromosome-based reporter, our HDR-USR system is a vector-based self-functional surrogate reporter coupled with a sgRNA and a donor vector for precision genome editing. The HDR-USR reporter vector contains four components: a *Cas9* gene expression cassette, a universal sgRNA expression cassette, a truncated puromycin-resistant gene ($Puro^R$) containing a universal sgRNA target sequence (sgT), and a $Puro^R$ gene sequence without an ATG and promoter ($\Delta Puro^R$) (used as an intra-molecular repair template). It is noteworthy that the universal sgT sequence specifically targets the HDR-USR but is not present at any site in mammalian genomes. When the HDR-USR reporter vector is transfected into cells, the expressed sgRNA guides the Cas9 protein to cleave its target sequence located in the middle of the truncated $Puro^R$ gene, resulting in a DSB in the vector. This DSB is then repaired by cellular DNA repair machinery. Two types of repaired vectors will be produced: NHEJ- and HDR-based repaired reporter vectors. Only the intra-molecular HDR-based repaired vector will express a functional $Puro^R$ gene (Figure 1A). Therefore, media supplemented with puromycin were used to enrich for cells that have experienced HDR-based repair events. Cells are simply co-transfected with the HDR-USR reporter vector, an sgRNA vector specifically targeting a desired chromosomal locus, and a homologous edited DNA donor vector. Transfected cells were grown in puromycin medium, and single clones were then picked for genotyping (Figure 1B).

To demonstrate its ability to enrich for HDR-based repaired cells, we tested our HDR-USR system to enrich for cells containing an HDR-mediated point mutation at the *EMX1* locus. To simplify detection, we mutated the protospacer-adjacent motif (PAM) of the *EMX1* target site to create an EcoRI recognition site on the homologous donor DNA sequence (Figure 1C). Then, HEK293T cells were co-transfected with pEMX1-sgRNA, pD-EMX1, and HDR-USR; selected with puromycin for 3, 5, 7, 10, and 15 days; and pooled for EcoRI digestion detection. Using HDR-USR enrichment, as shown in Figure 1D, the HDR-mediated cell pool enhanced 5.45- to 7.09-fold over the no-selection control with the puromycin selection time increase. To reduce random vector integration into chromosomes, we selected the shorter 5-day puromycin selection for the follow-up study.

We subsequently picked 50 cell clones from the HDR-USR group (selected by puromycin for 5 days) and 50 from the control group for genotyping via enzyme digestion and Sanger sequencing. Using EcoRI digestion, we found that only 2 clones (4%) were genetically modified in the control group, both of which were heterozygous (Figure S1). In striking contrast, we detected 31 clones (62%) modified with precision genome editing in the HDR-USR group, of which 23 were heterozygous, and 8 were homozygous (Figure S1). Sanger sequencing results revealed that 41 clones were wild-type (WT)/WT (no editing), 3 clones were WT/indel, 4 clones were indel/indel, and 2 clones were precision edited (PE)/indel in the control group. In contrast, the HDR-USR group contained 4 clones of WT/indel genotypes, 16 clones of indel/indel genotypes, 23 clones of PE/indel genotypes, and 8 clones of PE/PE genotypes, representing a marked enrichment in potentially useful clones (Figure 1E).

We next determined whether the HDR-USR system vector becomes randomly integrated into the genome or exhibits off-target effects. To detect random vector integration, we re-applied puromycin selection to positively cloned cells grown for 3 weeks without drug selection pressure. No clones were puromycin resistant (data not shown), suggesting that the puromycin-resistant cassette had not integrated.[27] Additionally, we designed four pairs of PCR primers to amplify different fragments of HDR-USR surrogate vector from PE cell clones in the HDR-USR group. No PCR product was detected for any clone, suggesting that random integration of the vector had not occurred in the selected clones (Figure S2A).

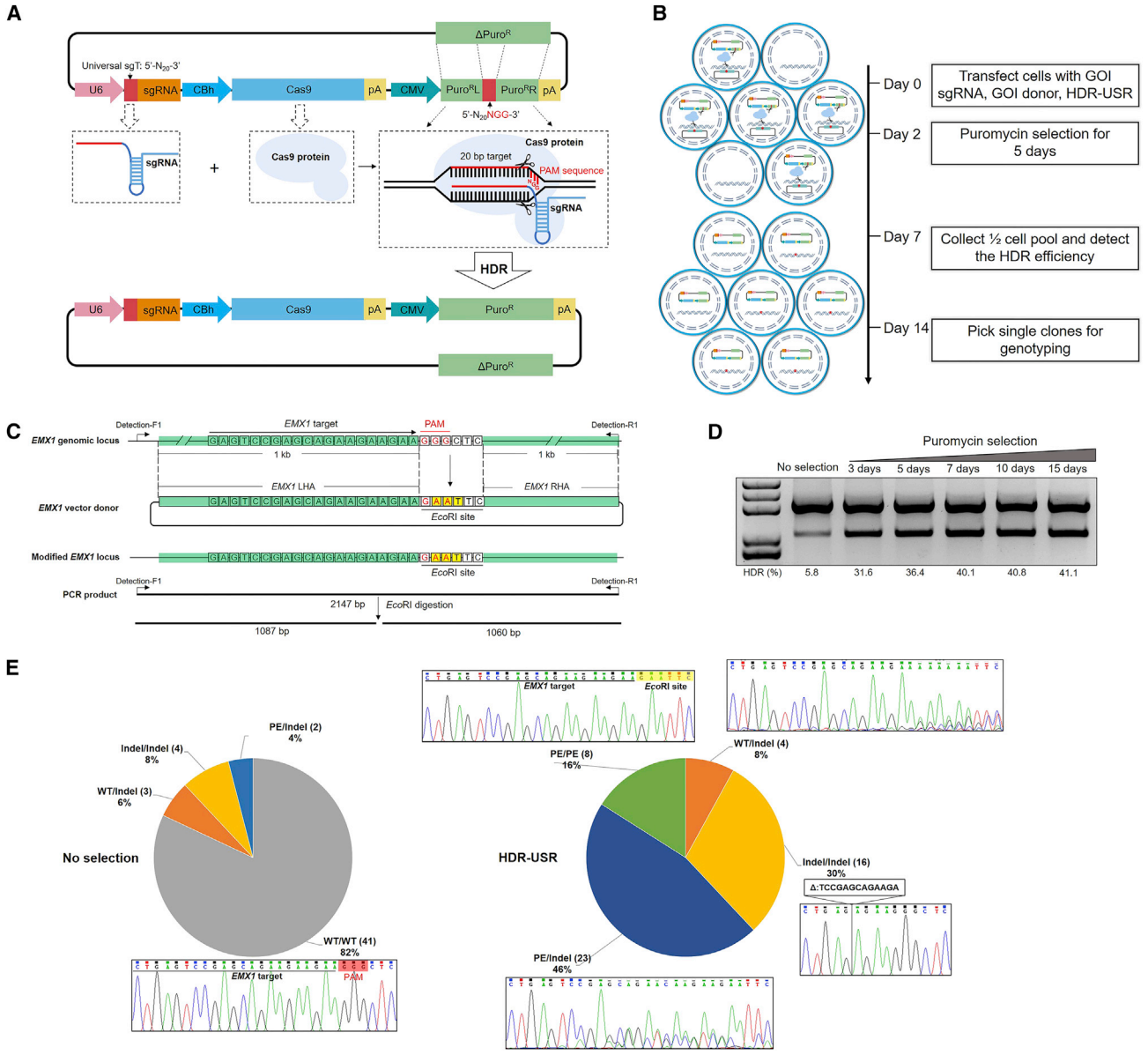


Figure 1. Function Principle of the HDR-USR and Proof of Feasibility

(A) Diagram of the HDR-USR system. The HDR-USR plasmid contains a CBh promoter-driven Cas9 expression cassette, a U6 promoter-driven universal sgRNA expression cassette, a CMV promoter-driven incomplete *Puro* sequence that deleted 100 bp and inserted a universal target sequence (5′-N_{20}NGG-3′), and a *Puro* sequence that does not contain the initiation codon ATG and promoter. The target sequence (5′-N_{20}-3′) of the universal sgRNA is present in the HDR-USR plasmid but not in the mammalian genome. (B) HDR-USR enrichment protocol for HDR-repaired cells. (C) Donor designs for point mutation of the *EMX1* locus. (D) Digestion assays of the cell pool at the *EMX1* locus in HEK293T cells with no selection or selected by puromycin for 3, 5, 7, 10, and 15 days. The HDR efficiency is indicated below each gel, which was calculated with the relative intensities of digested bands and the undigested band. (E) Genotypes of 50 clones in the no-selection group and 50 clones in the HDR-USR group at the *EMX1* locus in HEK293T cells. Representative Sanger-sequencing chromatograms are displayed for each genotype. WT, wild-type; PE, precision edited; Indel, insertion or deletion.

To examine whether the sgRNA encoded within the surrogate vector has any off-target effect on host cell chromosomes, we analyzed the top 10 potential off-target (OT) sites, as determined by https://zlab.bio/guide-design-resources, in the WT cells and the HDR-USR-enriched cells through deep sequencing. As a result, very low levels (<0.1%) of possible off-target mutations were detected at all 10 putative off-target sites, despite high levels of NHEJ efficiency at the *EMX1* locus (55.67%) (Figure S2B).

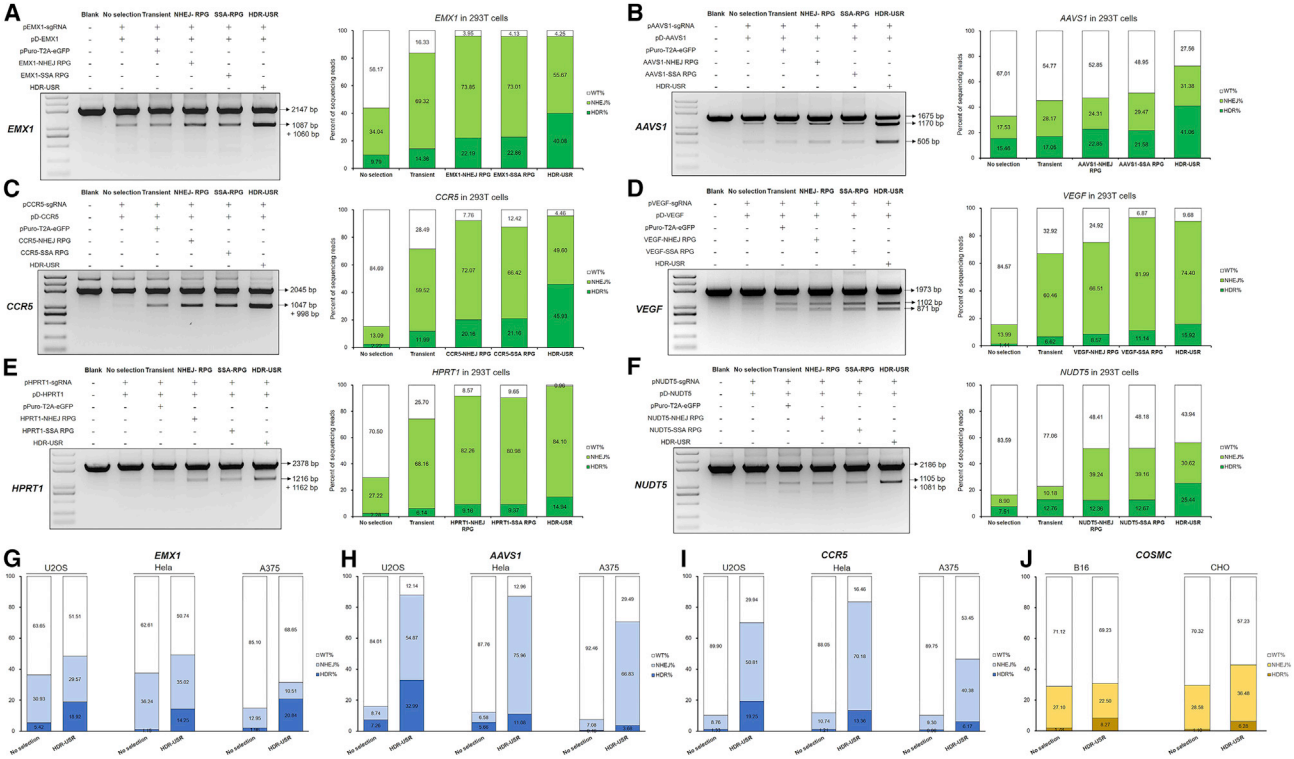


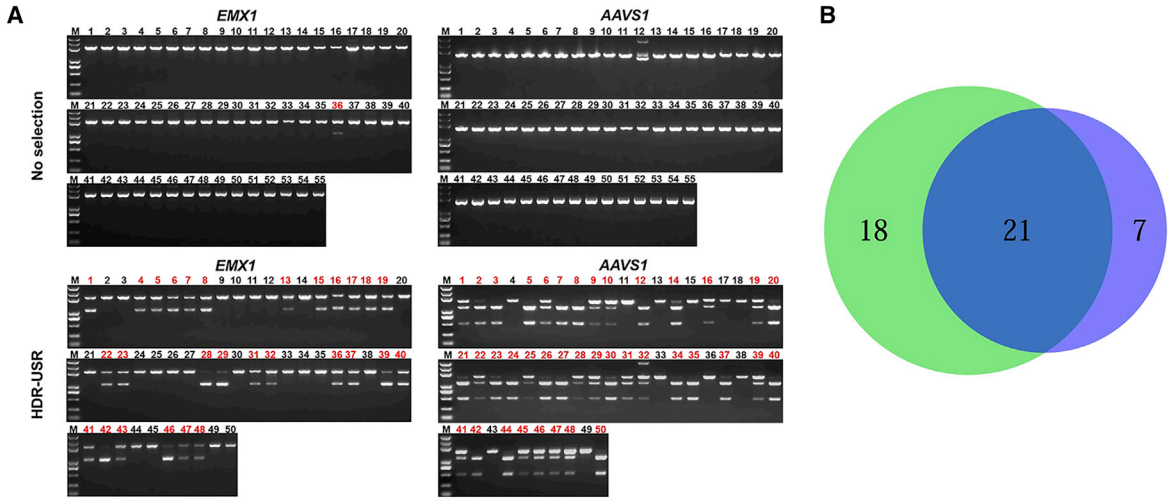
Figure 2. Efficient Enrichment of HDR-Mediated Precise Point Mutation at Various Loci in Multiple Cell Types

(A–F) Comparison of HDR-based precise point mutation efficiency with different surrogate reporters at (A) *EMX1*, (B) *AAVS1*, (C) *CCR5*, (D) *VEGF*, (E) *HPRT1*, and (F) *NUDT5* loci in HEK293T cells. (G–I) Comparison of HDR-based precise point mutation efficiency with or without HDR-USR plasmid in U2OS, HeLa, and A375 cells at (G) *EMX1*, (H) *AAVS1*, and (I) *CCR5* loci. (J) Comparison of HDR-based precise point mutation efficiency with or without HDR-USR plasmid in B16 and CHO cells at *COSMC* locus. Blank, cell group that did not transfect plasmid; No selection, cell group that co-transfected with Cas9 and donor plasmid; NHEJ-RPG, cell group that co-transfected with Cas9, donor, and the relative NHEJ-RPG plasmid; SSA-RPG, cell group that co-transfected with Cas9, donor, and the relative SSA-RPG plasmid; HDR-USR, cell group that co-transfected with Cas9, donor, and the HDR-USR plasmid. The results are indicated as the mean of three repeated transfections.

Collectively, these results indicate that the HDR-USR system can be used to efficiently enrich HDR-mediated precision point mutation at the *EMX1* locus in HEK293T cells, without detectable random vector integration events and with very low off-target effects.

Enrichment for Cells with Precision Point Mutations at Various Loci in Multiple Cell Lines

We next tested whether the HDR-USR system was equally efficacious at different loci and in different cell lines. First, we targeted six human genes—*EMX1*, *AAVS1*, *CCR5*, *VEGF*, *HPRT1*, and *NUDT5*—in HEK293T cells to create specific point mutations. To simplify detection, the PAM sequence of the sgRNA target sites within the donor was replaced by the restriction endonuclease sites, allowing detection of HDR-mediated precision genome editing by the digestion assay. In addition, to compare the efficiency of our HDR-USR system with previously reported surrogate reporters, we constructed three other reporter plasmids: pPuro-T2A-EGFP (Figure S3A) for enrichment of transfected cells and GOI (gene of interest)-NHEJ-RPG (DsRed-Puro^{R}-eGFP) (Figure S3B) or GOI-SSA-RPG (Figure S3C) for enrichment of nuclease-active cells.[29] 48 h after transfection, cells were subjected to either no selection (the control group) or puromycin selection for 5 days (the transient, NHEJ-RPG, SSA-RPG, and HDR-USR groups) (Figure S3D), and then the cells were harvested for detection. As shown in Figures 2A–2F, enzyme digestion assays clearly indicated that the cells with precision genome editing were most efficiently enriched by the HDR-USR system, followed by the NHEJ-RPG and SSA-RPG approaches. The pPuro-T2A-EGFP transient transfection selection method was the least effective. Cell-clone DNAs were analyzed via deep sequencing for validation. For the *EMX1* locus, the HDR-mediated precision-editing efficiencies were 9.79%, 14.36%, 22.19%, 22.86%, and 40.08% for the no-selection, transient, NHEJ-RPG, SSA-RPG, and HDR-USR groups, respectively. Thus, the HDR-mediated precision-editing efficiency in the HDR-USR group increased more than 4-fold over the no-selection group and about 2-fold compared to the other two surrogate reporter groups (Figure 2A). For the other five loci, we found that the HDR-mediated editing efficiency in the HDR-USR groups ranged from 14.94% to 41.06% and exhibited a 2.7- to 20.7-fold increase compared to the control. Most importantly, the average



C

Clone	EMX1	AAVS1	Clone	EMX1	AAVS1	Clone	EMX1	AAVS1	Clone	EMX1	AAVS1	Clone	EMX1	AAVS1
1	PE/Indel	PE/Indel	11	Indel/Indel	Indel/Indel	21	Indel/Indel	PE/PE	31	PE/Indel	PE/Indel	41	PE/Indel	PE/Indel
2	Indel/Indel	PE/PE	12	Indel/Indel	PE/PE	22	PE/Indel	PE/Indel	32	PE/Indel	PE/Indel	42	PE/PE	PE/PE
3	Indel/Indel	PE/PE	13	PE/Indel	Indel/Indel	23	PE/Indel	PE/PE	33	Indel/Indel	Indel/Indel	43	PE/Indel	Indel/Indel
4	PE/Indel	Indel/Indel	14	Indel/Indel	PE/PE	24	Indel/Indel	PE/PE	34	Indel/Indel	PE/PE	44	Indel/Indel	PE/PE
5	PE/Indel	PE/PE	15	PE/Indel	Indel/Indel	25	Indel/Indel	PE/Indel	35	Indel/Indel	PE/PE	45	Indel/Indel	PE/Indel
6	PE/Indel	PE/Indel	16	PE/Indel	PE/Indel	26	Indel/Indel	PE/PE	36	PE/Indel	Indel/Indel	46	PE/PE	PE/Indel
7	PE/Indel	PE/PE	17	PE/Indel	Indel/Indel	27	Indel/Indel	PE/PE	37	PE/Indel	PE/PE	47	PE/Indel	PE/Indel
8	PE/Indel	PE/PE	18	PE/Indel	Indel/Indel	28	PE/PE	PE/Indel	38	WT/WT	WT/WT	48	PE/Indel	PE/Indel
9	Indel/Indel	PE/Indel	19	PE/Indel	PE/Indel	29	PE/PE	PE/Indel	39	PE/PE	PE/Indel	49	Indel/Indel	Indel/Indel
10	Indel/Indel	PE/Indel	20	Indel/Indel	PE/PE	30	Indel/Indel	PE/PE	40	PE/Indel	PE/PE	50	Indel/Indel	PE/PE

Figure 3. HDR-USR Enriches for HDR-Mediated Dual Modification at *EMX1* and *AAVS1* Loci in HEK293T Cells

(A) Digestion genotyping at *EMX1* and *AAVS1* loci of the 55 cell clones in the no-selection group and 50 clones in the HDR-USR group. (B) Venn diagram of the cell clones for single or double modification at *EMX1* and *AAVS1* loci in the HDR-USR group. (C) Genotypes of 50 clones at *EMX1* and *AAVS1* loci in the HDR-USR group. WT, wild-type; PE, precision edited; Indel, insertion or deletion.

precision-editing efficiency across the five loci in the HDR-USR groups increased at least 2-fold compared to the other groups (Figures 2B–2F). Additionally, we selected another 10 target sites (*CUL3-site 1*, *CUL3-site 2*, *DNMT1*, *FANCF-site 1*, *FANCF-site 2*, *FANCF-site 3*, *GRIN2B*, *MECP2*, *UBE3A-site 1*, and *UBE3A-site 2*) to test the applicability of the HDR-USR system at more loci. As shown in Figure S4, the HDR-USR system enriched HDR efficiencies, which were enhanced 2.3- to 6.1-fold compared to the no-selection groups. Moreover, we tested the HDR-USR system in other human cell lines—including U2OS, HeLa, and A375 cells at the *EMX1*, *AAVS1*, and *CCR5* loci—and in rodent cell lines, including B16 and Chinese hamster ovary (CHO) at the *COSMC* locus. Due to lower transfection efficiencies in these lines, the absolute editing efficiencies (HDR and NHEJ) were comparatively lower than those in HEK293T cells. However, the HDR-mediated precision editing efficiencies in the HDR-USR group were all enhanced by 2.0- to 14.4-fold relative to the control group (Figures 2G–2J).

Together, these results validated the efficacy of our USR in enriching cells with HDR-mediated precision point mutations at various loci in multiple cell lines from several species.

Simultaneous Enrichment for Precision Editing at Two Loci

Given that editing multiple genes concurrently in gene function analysis and gene therapy is in high demand, we tested whether the HDR-USR system could be used for simultaneous enrichment for cells with double-locus precision genomic modifications. We co-transfected the HDR-USR surrogate reporter vector into HEK293T cells with two sgRNA expression vectors (pEMX1-sgRNA and pAAVS1-sgRNA) and two donor vectors (pD-EMX1 and pD-AAVS1). After puromycin selection, we picked 50 clones for genotyping via enzyme digestion and Sanger sequencing. As shown in Figure 3A, 21 of 50 clones (42%) were double edited at the *EMX1* and *AAVS1* loci. Among the 21 clones, only one clone was biallelically edited at both loci. For the *EMX1* locus, there were 28 clones (56%) with a PE heterozygous genotype, and 5 clones (10%) with a PE homozygous genotype. For the *AAVS1* locus, there were 39 clones (78%) with a PE heterozygous genotype and 21 clones (42%) with a PE homozygous genotype. In striking contrast, we did not detect any double-positive clones from 55 clones of the control group, and we found only one clone with heterozygous editing at the *EMX1* site (Figure 3). Thus, the HDR-USR system can simultaneously enrich for cells with precision modifications at multiple loci.

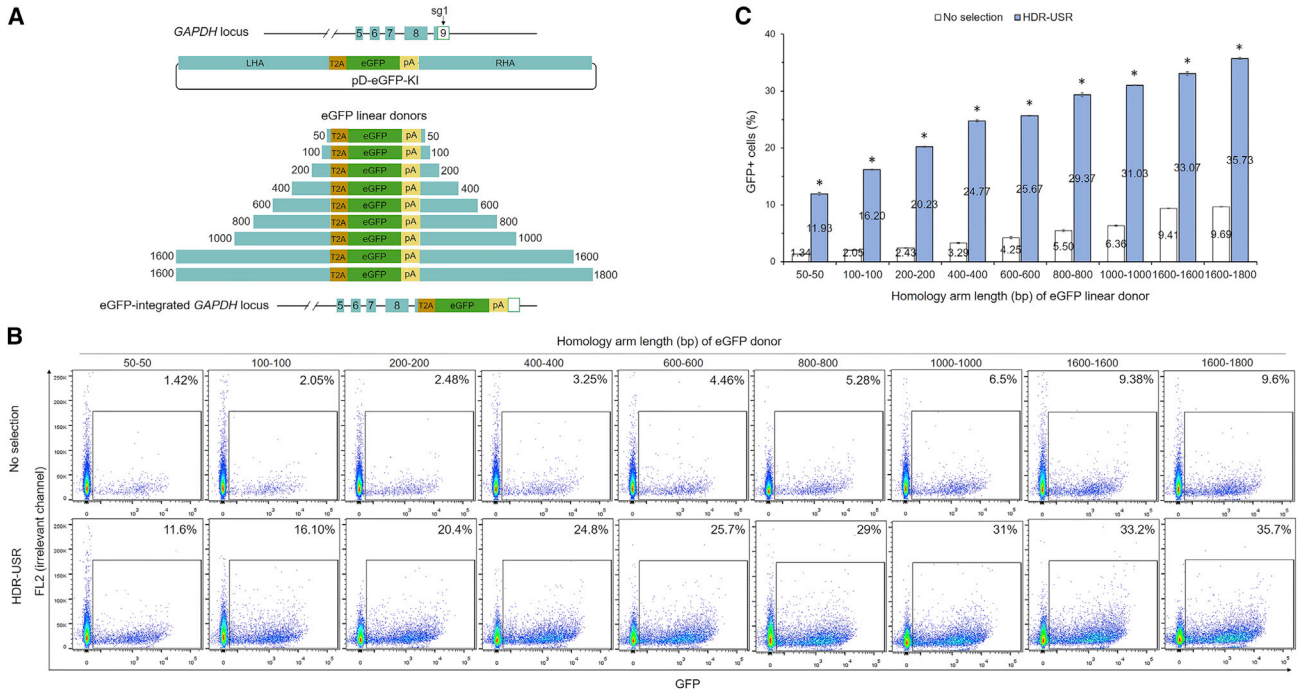


Figure 4. HDR-USR System Improves HDR-Mediated EGFP Knockin Efficiency

(A) Schematic outline of the pD-EGFP-KI plasmid donor and EGFP linear donors that were generated by PCR with HAs in the range of 50–1,800 bp in length. (B) Representative results of the flow-cytometric counting analysis for GFP$^+$ cells. (C) Statistical comparison of the percentages of GFP$^+$ cells in no-selection and HDR-USR groups. n = 3. Error bars indicate SEM. Asterisks indicate the significant difference between the HDR-USR and no-selection groups ($p < 0.05$).

Enrichment of HDR-Mediated Knockin Modifications

Precise knockin of foreign DNA into a selected genomic locus has opened up many exciting new possibilities for gene function studies and therapeutic genome editing. We next explored the application of the HDR-USR system for HDR-mediated knockin cells. To directly quantify and compare the efficiency of CRISPR/Cas9-induced HDR-mediated DNA integration, we constructed an *EGFP* knockin donor vector (pD-EGFP-KI), which has an *EGFP* sequence incorporated in frame prior to the *GAPDH* gene stop codon (Figure 4A; Figure S5A). This allows HDR-mediated knockin efficiency to be directly determined by flow cytometry (fluorescence-activated cell sorting [FACS]) analysis.[37]

To obtain the most effective sgRNA targeting of *GAPDH*, we first designed and constructed three sgRNA expression vectors—pGAPDH-sg1-Cas9 through pGAPDH-sg3-Cas9—and three corresponding SSA-RPG surrogate reporter vectors for quick targeting efficiency assay. The genome-targeting activities of three sgRNAs were compared, as we reported previously.[29,38] As demonstrated in Figure S5, *GAPDH*-sg1 showed the highest targeting activity and was used for subsequent experiments.

To examine the effect of HA length on HDR-mediated knockin efficiency, we generated a series of integration cassettes with different HA lengths ranging from 50 bp to 1,800 bp by PCR from the template plasmid pD-EGFP-KI (Figure 4A). These integration cassettes of different HA lengths were co-transfected into HEK293T cells with the HDR-USR surrogate vector and the sgRNA expression vector pGAPDH-sg1. FACS data analysis indicated that the 50-50-bp donor showed 1.34% knockin efficiency in the control group. Increasing HA length to 1,600–1,800 bp gradually raised efficiency to 9.69%. In contrast, the knockin efficiency of the 50-50-bp donor was 11.93% in the HDR-USR group. Increasing HA length from 100-100 bp to 1,600–1,800 bp increased knockin efficiency dramatically from 16.20% to 35.73%, which is equivalent to ~3.5- to 8.9-fold higher than in the control group (Figures 4B and 4C). Thus, with all the tested donors of different HAs, the HDR-USR system enables efficient enrichment for HDR-mediated knockin cells.

Enrichment of Biallelic, Large Deletion Modifications

The HDR mechanism can be also leveraged to generate a large DNA fragment deletion in the genome, which could be used to delete specific gene features or entire genes.[2] To examine the possibility that our HDR-USR can be applied to enrich for cells with large DNA fragment deletion, we co-transfected the HDR-USR vector into HEK293T cells with a DNA donor vector and two sgRNAs targeting the *AAVS1* locus. The deletion donor was designed to contain an *Xba*I restriction site at the junction (Figure 5A). After puromycin selection and cell expansion, we picked 36 clones for genotyping. We first looked for evidence of genome deletion by PCR analysis. The 2,899-bp band indicated amplification of the undeleted genome (including WT and single-indel disruption at one locus), and the 1,880-bp band

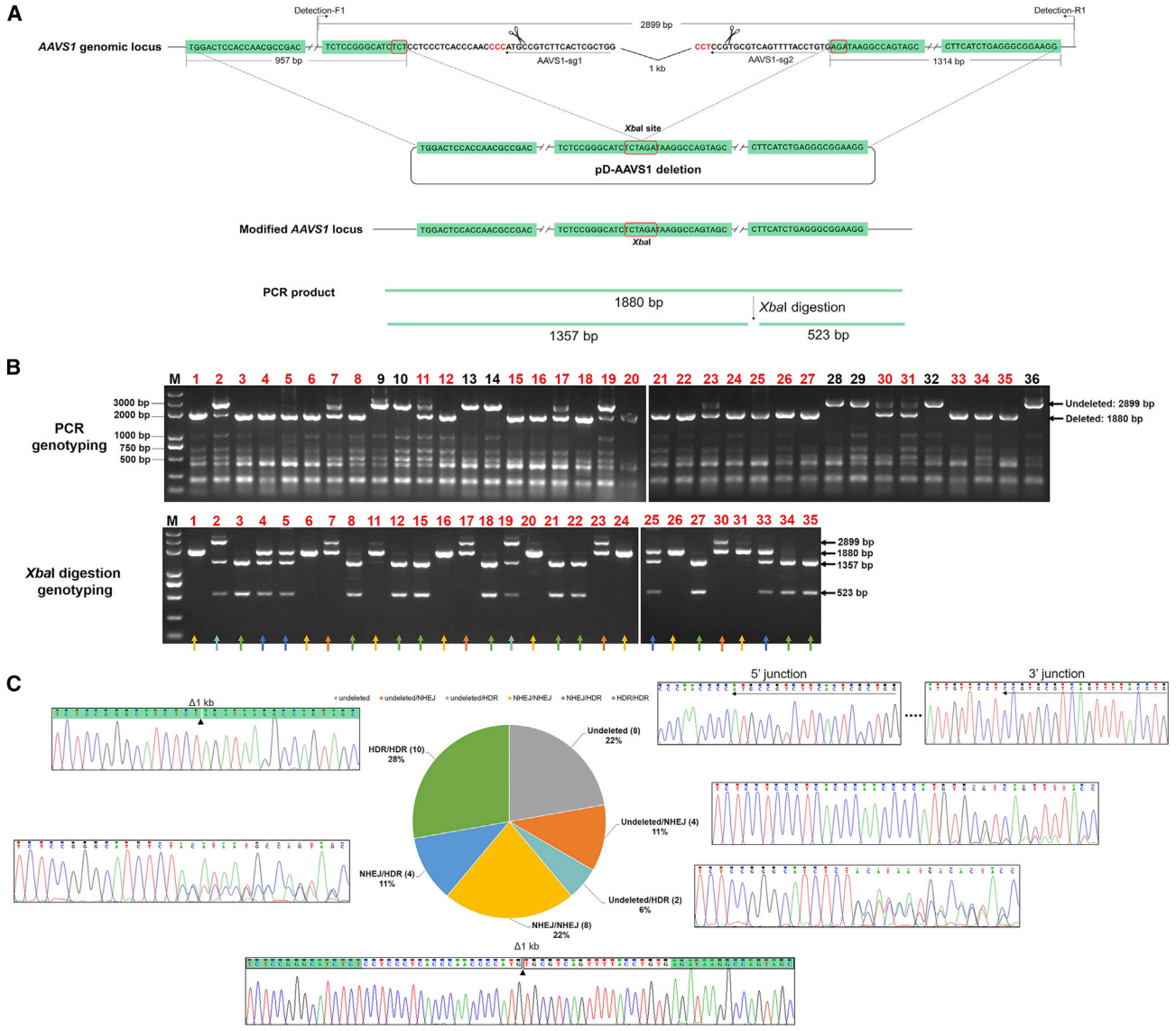


Figure 5. HDR-USR System Improves HDR-Mediated Precise Deletion

(A) Schematic illustration of the *AAVS1* genomic locus, paired sgRNAs (underlined with arrows), *AAVS1*-deletion donor, modified *AAVS1* locus, and *Xba*I digestion results after HDR deletion. The pD-AAVS1 deletion is designed to delete a 1-kb fragment at the *AAVS1* locus and create a precise junction (*Xba*I site) between the ends of the excised DNA. (B) PCR and *Xba*I digestion genotyping of the screened cell clones in the HDR-USR group in HEK293T cells. For PCR genotyping, the 2,899-bp band represents undeleted genome, and the 1,880-bp band represents the 1-kb deleted genome. For *Xba*I digestion genotyping, the 1,880-bp band represents NHEJ-mediated deletion, and the 1,357-bp and 523-bp bands represent HDR-mediated deletion. (C) Genotyping results of the 36 clones in the HDR-USR group by Sanger sequencing. Undeleted, WT and single-indel disruption at one locus; HDR, HDR-mediated precise deletion; NHEJ, NHEJ-mediated deletion.

represented amplification of the expected 1-kb deleted sequence. Among the 36 clones, 8 clones were undeleted, 6 clones exhibited heterozygous deletion, and the remaining 22 clones contained biallelic deletions (Figure 5B, PCR genotyping). The total deletion efficiency with HDR-USR enrichment reached up to 78% (28/36).

To distinguish HDR- and NHEJ-mediated deletions, we then digested PCR products with *Xba*I enzyme. As the donor fragment was designed to be incorporated with the *Xba*I recognition sequence, we considered the clone to represent an HDR-mediated editing allele if the 1,880-bp PCR product could be cut into 1,357-bp and 523-bp bands. Otherwise, the clone was classified as a NHEJ-mediated deletion. In addition, we further analyzed the clones by Sanger sequencing. Both *Xba*I digestion and sequencing results confirmed that, of the 28 deleted clones, 10 clones had homozygous HDR/HDR genotypes (35.71%), 8 clones had homozygous NHEJ/NHEJ

genotypes (28.57%), 4 clones had heterozygous NHEJ/HDR genotypes (14.29), 2 clones had heterozygous (undeleted/HDR-mediated) genotypes (7.14%), and 4 clones had heterozygous (undeleted/NHEJ-mediated) genotypes (14.29%) (Figures 5B and 5C). It is noteworthy that the percentage of precisely biallelic deletions enriched by HDR reached up to 28% (10/36), and the total percentage of biallelic deletion mediated by HDR or NHEJ was 61% (28% HDR/HDR + 22% NHEJ/NHEJ + 11% NHEJ/HDR).

In contrast, we picked 58 clones in the control group. PCR genotyping detected that only 5 clones displayed deleted genotypes at either one or both alleles (Figure S6A). *Xba*I digestion suggested that all the 5 deleted genotypes, including homozygote and heterozygotes, were induced by NHEJ and not by HDR-mediated mechanism (Figure S6B). As a result, the deletion efficiency in the control group without enrichment was 8.6%, and the percentage of biallelic deletion mediated by NHEJ was 1.7%.

Taken together, these results suggested that the HDR-USR system is a powerful tool for enriching for cells with biallelic HDR-mediated large DNA fragment deletions.

Further Improvement of HDR-USR Enrichment Efficiency

Although these experiments show that the HDR-USR system is an excellent and powerful tool for efficient enrichment of cells containing HDR-mediated point mutation, knockin, and deletion events, we pursued further optimization to maximize efficiency.

Given that Cas9 cuts the truncated reporter gene on both the HDR-USR vector and the genomic locus, we hypothesized that the reporter vector editing efficiency reflects the real efficiency of HDR-mediated genome editing on chromosomes. To explore the relationship between the cleavage of the surrogate reporter and chromosomal target, we designed three additional universal sgRNA targeting sequences (sgT1, sgT2, and sgT3) with different targeting efficiency scores predicated by the online software https://zlab.bio/guide-design-resources. To verify these predicted targeting efficiencies, we constructed the corresponding SSA-based surrogate reporter vectors and tested them with Cas9 in HEK293T cells by FACS, as previously reported.[29,38] The FACS data revealed the targeting efficiencies of the following sgRNA targeting sequences, in descending order as follows: sgT1, sgT, sgT2, and sgT3 (Figure S7), consistent with software prediction. Nevertheless, the differences of targeting efficiencies among sgT1, sgT, and sgT2 were little. Next, we incorporated these four sgRNA targeting sequences into our HDR-USR reporter vector and examined their respective capacities for enriching for cells with HDR-mediated editing. At the *EMX1* locus, the HDR-USR vector with sgT1 demonstrated the highest enrichment efficiency for HDR-mediated precision editing, while the HDR-USR with sgT3 showed the lowest enrichment efficiency. At the *AAVS1* and *HPRT1* loci, the HDR-USR with sgT3 also exhibited the lowest HDR-mediated chromosomal targeting efficiency, whereas the sgT group exhibited the highest chromosomal targeting efficiency. The sgT1 group showed a comparable efficiency with the sgT group (Figure 6A). Thus, the HDR-USR system using an sgRNA with high targeting efficiency was the most effective for the enrichment of the cells with the HDR-mediated genome precision editing.

Second, as small molecules or proteins that inhibit NHEJ can enhance the HDR repair pathway, and prolonging the S and G2 phases of the cell cycle increases HDR-mediated targeting efficiency,[10–15] we assessed whether these approaches could further increase the efficiency of HDR-USR-mediated precision editing. Specifically, we examined the effects of adenovirus 4 E1B55K and E4orf6 proteins (Ad4E1B-E4orf6)[13] and yeast Rad52 (yRad52)[11] by constructing vectors co-expressing these genes with Cas9 (Figure S8). Co-expression of yRad52 increased HDR targeting efficiency by 50%, 15%, and 260% at the *EMX1*, *AAVS1*, and *HPRT1* loci. Co-expression of Ad4E1B-E4orf6 increased HDR efficiency by 23%, 15%, and 57%, respectively (Figure 6B, left). The HDR editing efficiency with SCR7[10] exhibited a slight increase in editing at the *EMX1* and *HPRT1* sites but a modest decrease at the *AAVS1* locus. Nocodazole[15] did not enhance HDR-mediated genome editing events at any of the three loci (Figure 6B, right). Collectively, co-expression of yRad52 or Ad4E1B-E4orf6 with our HDR-USR system yielded enhancement of HDR-mediated precision genome editing efficiency.

Finally, we compared different types of donors, including a conventional plasmid donor, an ssODN donor, a linear donor, a double cut plasmid donor,[24] and a double cut linear donor. Including the Cas9 target site(s) on either linear or plasmid donors has also been shown to increase integration efficiency.[24] As shown in Figure 6C, the HDR-mediated editing efficiencies with different concentrations of the 89-nt ssODN donor were all lower than those with conventional plasmid donors whose left HA (LHA) and right HA (RHA) were about 1 kb long. Notably, increasing the HA length of the ssODN (>110 nt) enhanced HDR efficiency when the concentrations were more than 500 ng, resulting in a 1.47-fold enhancement with 5 μg of 130 nt ssODN donor. In addition, linearized plasmid donor and double cut donor significantly enhanced our HDR-USR enrichment efficiency for HDR-mediated precision genome editing, on average, by 1.25- to 2.67-fold at three loci (Figure 6D).

Taken together, these data suggest that the enriched precise editing with our HDR-USR system can be further improved. Using the HDR-USR system with high sgRNA activity, we can achieve coexpression of yRad52 or Ad4E1B-E4orf6 and use ssODN donors longer than 110 nt or a linear double-stranded DNA donor to efficiently enrich the HDR edits.

DISCUSSION

In the present study, we developed the HDR-USR system for enrichment of cells with HDR-mediated precision genome editing. As distinguished from the previous "co-targeting with selection" methods,[34–36] our system functions by itself in an episomal manner. A key component of the system is a surrogate reporter containing a truncated puromycin-resistant gene. Puromycin resistance is restored

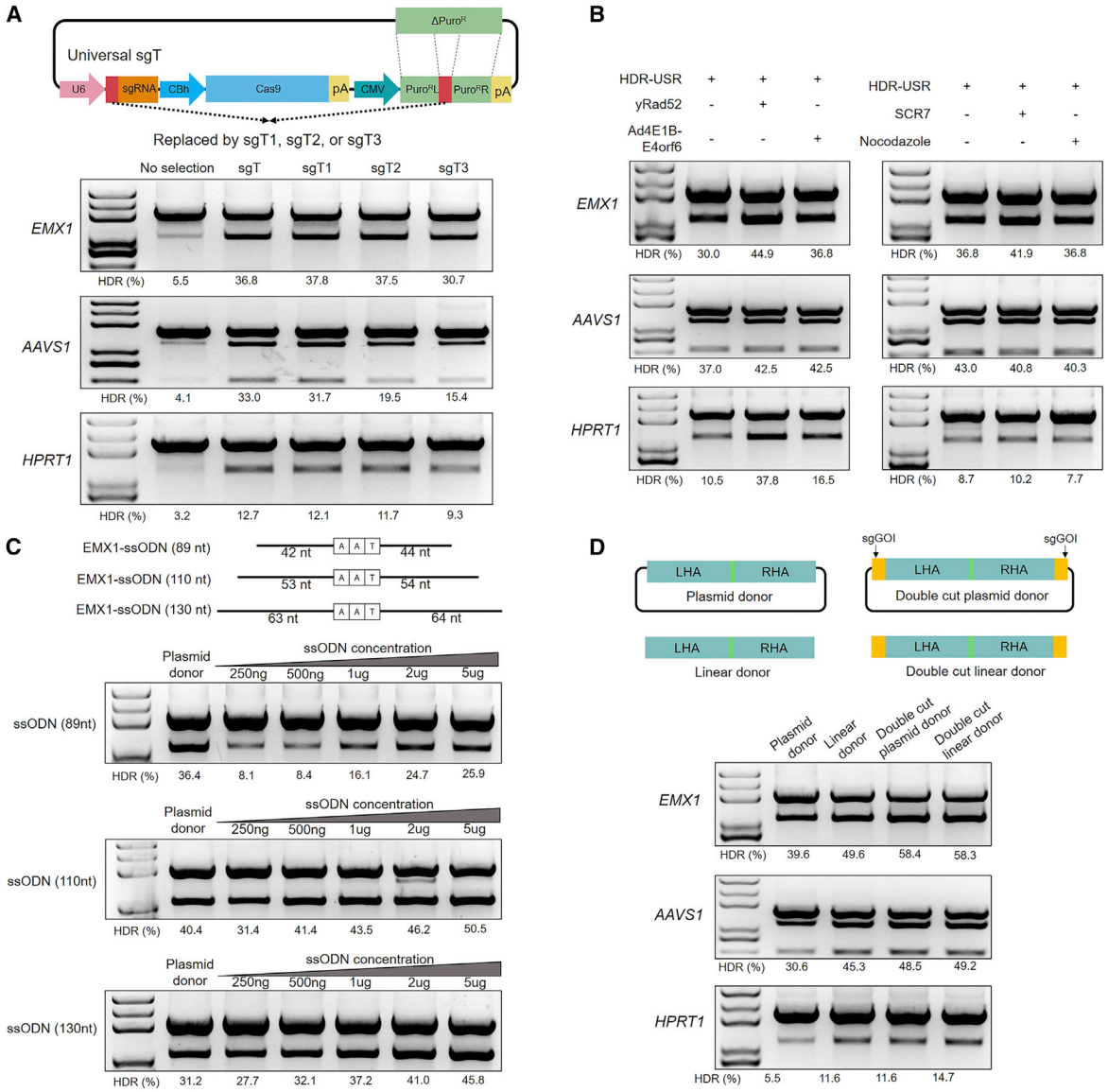


Figure 6. Improved HDR-Based Precise Genome Editing Efficiency with the HDR-USR System

(A) The effects of different target sequences in the HDR-USR plasmid on HDR efficiency at *EMX1*, *AAVS1*, and *HPRT1* loci in HEK293T cells. (B) The effects of yRad52, Ad4E1B-E4orf6, SCR7, and nocodazole on HDR efficiency at *EMX1*, *AAVS1*, and *HPRT1* loci in HEK293T cells. (C) The effects of ssODN donors (with 89-nt, 110-nt, and 130-nt lengths) on HDR efficiency at the *EMX1* locus in HEK293T cells. (D) The effects of the plasmid donor, linear donor, double cut plasmid donor, and double cut linear donor on HDR efficiency at *EMX1*, *AAVS1*, and *HPRT1* loci in HEK293T cells.

by self-cleavage and self-repair from expressed Cas9 protein guided by a specific sgRNA and a homologous intra-molecular template via HDR. Thus, the puromycin expression cassette can be used to enrich for cells that have undergone HDR-mediated precision genome editing. The HDR-USR system is simple, broadly applicable, and effective at enhancing HDR-based precision genome editing (including point mutation, knockin, and deletion) across different loci and in various cell types and can be further improved by expression of yRad52 with a linear single-/double-stranded DNA donor.

Point mutation is the most common genetic mutation known to be responsible for many inherited disorders. Recently, cytosine base editors (CBEs) and adenine base editors (ABEs), which can efficiently install the conversion of C to T and A to G, were widely used in mammalian cells,[39,40] but they cannot catalyze transversion mutations currently. However, traditional HDR can convert any base to any other base, even though the efficiency is low. In this study, we initially focused on point mutation to test our HDR-USR system, and 31 of 50 (62%) PE clones were obtained at the *EMX1* locus in

HEK293T cells. The HDR efficiency enriched by HDR-USR is comparable with that enriched by co-targeting methods in human induced pluripotent stem cells (iPSCs) (~40%)[34] or in mouse embryonic stem cells (ESCs) (~65%).[35] These results suggest that the HDR of DSBs on vectors and genomes may be carried out simultaneously by the same mechanisms.

Vector-based surrogate reporters have often been used to enrich for either transfected cells or functional artificial nucleases with puromycin antibiotic.[31] With regard to the enrichment efficiency, we systematically compared our HDR-USR system with pPuro-T2A-EGFP, which is used for enrichment of transfection-positive cells, and NHEJ-RPG and SSA-RPG, which are designed for enrichment of nuclease-active cells[29] (Figure 2; Figure S3). Our HDR-USR system demonstrated extraordinary performance for enrichment of cells exhibiting HDR-mediated precision genome editing. At all six loci tested, the HDR-USR group achieved about 2-fold more enrichment efficiency than that of the NHEJ-RPG and SSA-RPG groups and up to 20-fold more than that of the control group. Nevertheless, NHEJ-RPG and SSA-RPG surrogate reporter systems may remain more suitable for enrichment of NHEJ-mediated editing cells. The enrichment efficiency of pPuro-T2A-EGFP for either HDR- or NHEJ-mediated editing cells was the lowest of the approaches tested.

Since complex diseases are often caused by multiple factors, correction of mutations at multiple sites could be necessary. Chromosome integration-based "co-targeting with selection" methods have the potential to enrich for simultaneous dual-locus modifications, and a previous report isolated 6 clones bearing both editing among selected 50 clones (12%).[34] In this study, we used the HDR-USR system to generate precise modifications at both the *EMX1* and *AAVS1* loci in 21 of 50 (42%) clones. This high efficiency of double loci editing suggests that our HDR-USR system will be useful for the enrichment of edited cells with targeted mutations at multiple loci.

HDR-mediated knockin or replacement of exogenous DNA at a selected genomic locus is anticipated to facilitate gene function study and develop novel gene therapy strategies. For efficient HDR-based knockin, HA length is the key consideration for donor template design. We observed significantly enhanced knockin efficiency with all the tested linear donors with HA lengths ranging from 50 to 1,800 bp using the HDR-USR system. As expected, increasing HA length enhanced the HDR efficiency as previously reported.[2] Notably, the donors with HA lengths longer than 1,000-1,000 bp exhibited relatively higher HDR efficiency. Thus, for efficient enrichment of knockin of large fragments with the HDR-USR system, the HA should be at least 1,000 bp.

A single sgRNA in the CRISPR/Cas9 system has been often used to induce indels at a chromosome locus. If the sgRNA target site is located within a gene exon, these indels most likely lead to frameshift/premature termination or stop codon mutations, thereby inactivating the corresponding protein. The biggest disadvantage of this approach to knock out a gene comes from its uncontrolled indels. Two sgRNAs coupled with a carefully designed homologous donor in the CRISPR/Cas9 system can be leveraged to generate specific and controllable deletions of target chromosomal sequences. This technology provides a powerful tool to study specific gene features, intergenic or intronic regulatory sequences, or noncoding RNA genes.[41–45] Thus far, there are very few reports on using a surrogate reporter for the enrichment of cells with scarless, large deletions of target genomic sequences. In this application, our HDR-USR system achieved up to 78% deletion efficiency induced by either HDR or NHEJ, with 79% of those clones exhibiting biallelic deletion. Interestingly, NHEJ-induced deletion genotypes largely resulted from precise ligation of the two blunt-ended DSBs produced by Cas9, with each DSB occurring at exactly 3 bp upstream of the PAM sequence (Figure 5). This phenomenon was also observed previously.[42–44,46] Understanding the mechanism leading to these precise ligations may help to harness accurate NHEJ to generate precise indels of defined length. Collectively, our HDR-USR system proved to be an excellent surrogate reporter for enrichment of the cells with CRISPR/Cas9-mediated precision deletion.

Previous surrogate reporter systems incorporate the genome targeting sequence into the middle of a reporter gene, thereby reporting nuclease activity.[29–31] Thus, whether the nuclease activity targeting the genomic loci was influenced by targeting reporters has not been accurately assessed. In this study, we introduced a universal sgRNA target site sequence, which is not derived from either the human genome or the mouse genome, into the middle of a truncated puromycin-resistant gene. As such, the HDR-USR surrogate reporter works independently. This design allows us to estimate the relationship between Cas9 cleavage activity at the surrogate reporter and the genomic loci. We designed four universal sgRNA target sites with different targeting activities and verified their targeting activities using SSA-based surrogate reporter experiments. We incorporated these four universal sgRNA targeting sequences into our HDR-USR surrogate reporter vector and found that a surrogate reporter with a relatively high sgRNA activity could exert highly efficient HDR enrichment (Figure 6A), implying that the target activity in surrogate reporters may positively correlate with the enrichment efficiency.

Since the HDR-USR system was designed to enrich for cells with HDR-mediated precision genome editing, we reasoned that HDR boosters might further improve the enrichment efficiency of the HDR-USR system. Ad4E1B-E4orf6, which degrades ligase IV, was reported to promote HDR efficiency 8-fold.[13] yRad52, which is involved in the HDR process, can improve the efficiency 3-fold.[11] Small molecules like SCR7[47,48] and nocodazole[15,16] also increase HDR events. After carefully examining these factors, we found that co-expression of yRad52 with the HDR-USR system provided the greatest degree of enhancement, with a 3.6-fold increase. This is consistent with a previous observation.[11] Neither SCR7 nor nocodazole had a significant positive effect on efficiency (Figure 6B). Notably, the mortality rates for cells coexpressing Ad4E1B-E4orf6 and for the nocodazole-treated group were higher than those in the other groups. We cannot rule out

the possibility that expression of both adenoviruses E1B55K and E4orf6 could have toxic effects on HEK293T cells. Functioning as synchronizing cell cycle at the G2/M phase, nocodazole might interfere cell resistance to puromycin.

As the second part of our HDR-USR system, the donor vector plays an important role in high efficiency enrichment of cells with HDR-mediated precision genome editing. Traditionally, ssODN donors are used for genome base change editing, and plasmid donors are often used for larger genome changes.[2,7] Linear donors generated by PCR amplification and double cut donors flanked by sgRNA-targeting sequences are efficient for genome editing in human cells.[24,49] We comprehensively compared different donor structures for enriched-HDR efficiency with our HDR-USR system. For point mutations at three loci, we showed ssODN donors longer than 110 nt, linear donors, and double cut donors (plasmid or linear) were all superior to plasmid donors whose HAs were 1 kb. Combined with previous conclusions, when enriched with our HDR-USR system, we recommend the use of ssODN donors (>110 nt) for small editing and linear donors or double cut donors (HAs > 1 kb) for knockin or deletion of large fragments.

In summary, we devised an HDR-USR system for efficient enrichment of the cells with CRISPR/Cas9-induced HDR-mediated precision genome editing. This system provides a simple, robust, and efficient surrogate reporter for enrichment of the cells with CRISPR/Cas9-induced HDR-mediated precision genome editing without perturbation of genome integrity, as it works in an episomal manner. The HDR-USR surrogate reporter vector works universally for any purpose commonly required for precision genome editing projects with a highly efficient enrichment capacity.

MATERIALS AND METHODS

Construction of the HDR-USR Vectors

Three main steps were needed to construct the HDR-USR reporter vector. First, pXL-CMV-Puro^{R}L-sgT_NGG-Puro^{R}R-polyA (polyadenosine) was constructed. Specifically, Puro^{R}L-sgT_NGG-Puro^{R}R was amplified by overlap PCR and inserted into the transitional vector pcDNA3.1(+) to generate pcDNA3.1-CMV-Puro^{R}L-sgT_NGG-Puro^{R}R-polyA, and then the CMV-Puro^{R}L-sgT_NGG-Puro^{R}R-polyA cassette was amplified from the transitional vector and inserted into the pXL-BACII backbone. Second, the puromycin gene sequence, which does not contain the initiation codon ATG (ΔPuro) was amplified and inserted into the plasmid from the first step to construct pXL-CMV-Puro^{R}L-sgT_NGG-Puro^{R}R-polyA-ΔPuro. Third, pXL-U6-sgT-CBh-Cas9-polyA-CMV-Puro^{R}L-sgT_NGG-Puro^{R}R-polyA-ΔPuro (namely, HDR-USR) was constructed. In particular, reverse complementary oligonucleotide pairs of 20 bp corresponding to the universal sgT target sequence with *Bbs*I overhangs were synthesized, annealed, and cloned into pX330-U6-Chimeric_BB-CBh-hSpCas9 (Addgene plasmid #42230)[3] to build psgT-Cas9; then, U6-sgT-CBh-Cas9-polyA was amplified from psgT-Cas9 and inserted into the plasmid from the second step to construct HDR-USR (Figure 1A). The other HDR-USR plasmids using different universal sgRNA target sequences were constructed similarly.

All the primers used for constructing the HDR-USR vectors are shown in Table S1. All constructs were confirmed by Sanger sequencing.

Construction of Transient Transfection Screening Vector, SSA-RPG and NHEJ-RPG Surrogate Reporter Vectors, and the sgRNA Vectors

For the construction of transient transfection screening vector pPuro-T2A-EGFP (Figure S3A), the Puro^R sequence was amplified and inserted into pRS426-CMV-T2A-EGFP-polyA[50] to construct pRS426-CMV-Puro-T2A-EGFP-polyA; namely, pPuro-T2A-EGFP. The NHEJ-RPG and SSA-RPG vectors (Figures S3B and S3C) for *EMX1*, *AAVS1*, *CCR5*, *VEGF*, *HPRT1*, *NUDT5*, *GAPDH*-sg1–*GAPDH*-sg3, and universal sgT/T1/T2/T3 were constructed as previously reported.[29]

sgRNA vectors for different targets (Table S2) included pGOI-sgRNA and pGOI-sgRNA-Cas9. pGOI-sgRNA was used for co-transfection with HDR-USR, which contains a Cas9 gene. pGOI-sgRNA-Cas9 was used for co-transfection with the other three surrogate reporter vectors. For the construction of pGOI-sgRNA-Cas9, reverse complementary oligonucleotide pairs of 20 bp corresponding to GOI-target sequences with *Bbs*I overhangs were synthesized, annealed, and cloned into pX330-U6-Chimeric_BB-CBh-hSpCas9. For the construction of pGOI-sgRNA, U6-GOI-sgRNA was amplified from pGOI-sgRNA-Cas9 and cloned into pBlueScript II SK(+).

All primers used for constructing the surrogate reporter vectors and the sgRNA vectors are shown in Table S3.

Construction of the Donor Plasmids

The donor plasmids used in this study included three types: donors for point mutations (including conventional plasmid donors and double cut plasmid donors), donors for fragment knockin, and donors for precise deletion.

Conventional plasmid donor vectors for point mutations contained pD-EMX1, pD-AAVS1, pD-CCR5, pD-VEGF, pD-HPRT1, pD-NUDT5, pD-CUL3 site 1, pD-CUL3 site 2, pD-DNMT1, pD-FANCF site 1, pD-FANCF site 2, pD-FANCF site 3, pD-GRIN2B, pD-MECP2, pD-UBE3A site 1, pD-UBE3A site 2, pD-mCOSMC, and pD-CHO-COSMC. The donor fragments were amplified by overlap PCR from WT genomic DNA (gDNA) of HEK293T cells, B16 cells, or CHO cells and cloned into pXL-BACII to generate the corresponding donor vectors. The LHA and RHA of the point mutation donors were set at about 1 kb. The PAM motifs of the sgRNA target sites within the donor constructs were replaced by the corresponding restriction endonuclease sites for digestion assays to detect HDR efficiencies. For building the double cut donor vectors pD-EMX1-sg, pD-AAVS1-sg, and pD-HPRT1-sg, both the forward and the reverse primers, which contained the corresponding sgRNA target sequence

together with a PAM (NGG) were used to amplify donor fragments from the constructed plasmids pD-EMX1, pD-AAVS1, and pD-HPRT1, respectively. The amplified donor templates harboring sgRNA recognition sites were subsequently cloned into pXL-BACII to generate double cut plasmid donor vectors, respectively.

For construction of *EGFP* knockin donor vector pD-EGFP-KI, T2A-EGFP-pA was first digested from pRS426-CMV-Puro-T2A-EGFP-polyA[50] and inserted into pXL-BACII to construct pXL-T2A-EGFP-pA. Then GAPDH LHA (about 1.6 kb) and GAPDH RHA (about 1.8 kb) were amplified separately from WT gDNA of HEK293T cells and inserted into pXL-T2A-EGFP-pA to construct pXL-GAPDH LHA-T2A-EGFP-pA-GAPDH RHA (namely, pD-EGFP-KI).

For construction of the *AAVS1*-deletion donor, the repair template sequence, which contained a *Xba*I restriction site at the junction, was amplified via overlap PCR from WT gDNA of HEK293T cells and cloned into pXL-BACII to create pD-AAVS1 deletion.

All the primers used for constructing the donor vectors are shown in Table S4.

Construction of Putative HDR-Enhancing Plasmids

To test whether Ad4E1B-E4orf6 and yRad52 could further enhance HDR efficiency after HDR-USR enrichment, pEMX1/AAVS1/HPRT1-sgRNA-Cas9-T2A-Ad4E1B-P2A-Ad4E4orf6 and pEMX1/AAVS1/HPRT1-sgRNA-Cas9-T2A-yRad52 were constructed. The detailed construction process was as follows: Ad4E1B was amplified from pU6-sgRosa26-1-CBh-Cas9-T2A-BFP-P2A-Ad4E1B (Addgene plasmid #64219) and inserted into pU6-(*Bbs*I)-CBh-Cas9-T2A-mcherry-P2A-Ad4E4orf6 (Addgene plasmid #64222)[13] to construct the intermediate vector pU6-sgRNA-CBh-Cas9-T2A-Ad4E1B-P2A-Ad4E4orf6. Then, T2A-Ad4E1B-P2A-Ad4E4orf6 was digested by *Fse*I and EcoRI from the intermediate vector and inserted into pEMX1/AAVS1/HPRT1-sgRNA-Cas9 to create pEMX1/AAVS1/HPRT1 -sgRNA-T2A-Ad4E1B-P2A-Ad4E4orf6. yRad52 was amplified from pCBh-Rad52[11] and inserted into pEMX1/AAVS1/HPRT1-sgRNA-T2A-Ad4E1B-P2A-Ad4E4orf6 to create pEMX1/AAVS1/HPRT1-sgRNA-T2A-Rad52.

All the primers used for constructing the two HDR-enhancing plasmids are shown in Table S5. Schematic diagrams of these plasmid are shown in Figure S8.

Cell Culture and Transfection

Human cell lines, including HEK293T, U2OS, HeLa, and A375 cells, and rodent cells including B16 and CHO cells were maintained in Dulbecco's modified Eagle's medium (DMEM; GIBCO) supplemented with 10% fetal bovine serum (FBS; ScienCell), 100 U/mL penicillin, and 100 μg/mL streptomycin in a 37°C humidified atmosphere with 5% CO_2 incubation.

Cells were seeded into 6-well plates 1 day before transfection. At 60%–70% confluency, cells were transfected using Lipofectamine 2000 Reagent (Life Technologies) according to the manufacturer's protocol. For each well, a total of 5 μg plasmids was used, and the molar ratio was 1:1:1. Three parallel transfections were performed for each group.

Detection of HDR-Based Precision Point Mutation Efficiency at Various Loci with Different Surrogate Reporters

Cells were co-transfected with pGOI-sgRNA-Cas9/pGOI-sgRNA and pD-GOI, along with one of the four surrogate reporters—pPuro-T2A-EGFP, GOI-NHEJ-RPG, GOI-SSA-RPG, or HDR-USR—using Lipofectamine 2000 Reagent (Life Technologies). After 48 h, cells were selected with puromycin (3 μg/mL for HEK293T cells; 1 μg/mL for U2OS, HeLa, and A375 cells; 2 μg/mL for B16 cells; and 7 μg/mL for CHO cells) for 5 days, and puromycin-resistant cells were harvested and pooled. Cells co-transfected with pGOI-sgRNA-Cas9 and pD-GOI were used as the "no-selection group" and were harvested after transfection for 48 h. The collected cells from different treatment groups were used for gDNA extraction for digestion assays and deep-sequencing analysis.

For digestion assays, the target regions of different loci were amplified by corresponding detection primers (shown in Table S6), purified, and digested with the corresponding restriction endonucleases. The proportion of the digested DNA was calculated by gray analysis to estimate HDR efficiency. The digestion results were collected from a pool of three repeated transfections.

For deep sequencing, the aforementioned purified PCR products from different target loci served as a template for the amplification of PCR amplicons by primers that added distinguishable barcodes. Amplicons were sequenced on an Illumina HiSeq X Ten (GENEWIZ). Among the FASTQ files, reads with the donor sequence (typically ~26 nt, spanning the targeted nucleotide and the mutant restriction enzyme cutting site) were considered as HDR-mediated editing. Reads corresponding to the NCBI reference sequence were regarded as WT, and the sum of reads subtracting WT reads and HDR reads were considered as reads corresponding to NHEJ-mediated editing.

Genotypic Analysis of Cell Clones from Point Mutations at the *EMX1* Locus, Simultaneous Dual Editing at the *EMX1* and *AAVS1* Loci, and Precise Deletion at the *AAVS1* Locus Using the HDR-USR System

Cells co-transfected with pGOI-sgRNA-Cas9 and pD-GOI were used as controls, and cells co-transfected with pGOI-sgRNA, pD-GOI, and HDR-USR were used as the experimental group. Both the control cell groups which transfected for 48 h and the experimental cell groups which transfected and selected with puromycin for 5 days were diluted and seeded into 96-well plates. 3 days later, single-cell clones were evaluated to exclude multiple cell contamination events. Cell clones were cultured until confluence and were then transferred into 24-well plates for proliferation. 30–60 cell clones in each group were collected for analysis. Half of the cells from each clone were used for gDNA extraction and PCR amplification by gene-specific detection primers (Table S6). Resulting amplicons were purified for enzyme digestion and Sanger sequencing.

Random Integration Detection and Off-Target Analysis

To detect the random integration of the HDR-USR vector in chromosomes, the retained positive cell clones for the target *EMX1* locus were maintained without drug selection pressure for 3 weeks. Puromycin was subsequently added to the medium, and the cells were observed after 3 days. In addition, four pairs of primers spanning the HDR-USR vector (Figure S2A; Table S7) were designed to amplify potential integration events from the gDNA of the PE clones at the *EMX1* locus.

For off-target analysis of the HDR-USR vector in cell chromosomes, 10 potential off-target sites for universal sgT sequence (GCGAATGC CACAAGCGGAGA AGG) were predicted based on the online software (https://zlab.bio/guide-design-resources). gDNA from WT or HDR-USR-enriched cells for precision editing of the *EMX1* locus were amplified using primers that added distinguishable barcodes. Amplicons were sequenced on an Illumina HiSeq X Ten (GENEWIZ). The indel frequencies were analyzed as previously reported.[51]

Flow Cytometry

For comparing the on-target activity of different sgRNAs, HEK293T cells were co-transfected with pGOI-sgRNA-Cas9 and GOI-SSA-RPG in 6-well plates and analyzed with RFP and GFP signals using a BD FACSAria III flow cytometer. As previously reported by our lab and as shown in Figure S3C, the RFP^+ cells represent the transfected cells, and GFP^+ cells stand for SSA-repaired cells. The percentage of RFP^+GFP^+-positive cells compared with the sum of RFP^+GFP^+ and RFP^+ cells was calculated as an indirect measurement to evaluate the sgRNA activities.[29,38]

To evaluate HDR-mediated knockin efficiency, HEK293T cells of the control groups (which co-transfected with pGAPDH-sg1-Cas9 and *EGFP* donors with different HA lengths and were harvested after 48 h) and those of the HDR-USR groups (which were co-transfected with pGAPDH-sg1-Cas9, *EGFP* donors, and HDR-USR and were harvested after selection with puromycin for 5 days) were analyzed by FACS. The percentage of GFP^+-positive cells was calculated as a direct assessment of the knockin efficiency.

Detection of the Effects of Different Universal sgRNA Targeting Sequences on HDR Efficiency

To test whether different sgRNA targeting activities in HDR-USR vectors could exert different effects on HDR efficiency, HEK293T cells were co-transfected with pEMX1/AAVS1/HPRT1-sgRNA; pD-EMX1/AAVS1/HPRT1; and HDR-USR, including sgT, sgT1, sgT2, or sgT3 (predicted scores were 88, 96, 73, and 52 by https://zlab.bio/guide-design-resources, respectively). After selection with puromycin for 5 days, cells were harvested for gDNA extraction and digestion assays.

Detection of the Effects of yRad52, Ad4E1B-E4orf6, SCR7, and Nocodazole on the Enrichment Efficiency of the HDR-USR System

To detect whether co-expression of putative HDR-enhancing genes yRad52 or Ad4E1B-E4orf6 could further improve the enrichment efficiency of the HDR-USR system, HEK293T cells were co-transfected with pD-EMX1/AAVS1/HPRT1, HDR-USR vector, and pEMX1/AAVS1/HPRT1-sgRNA-Cas9-T2A-yRad52 (or pEMX1/AAVS1/HPRT1-sgRNA-Cas9-T2A-Ad4E1B-P2A-Ad4E4orf6). After transfection for 48 h, cells were selected with puromycin for 5 days and then harvested for gDNA extraction and digestion assays.

To test the effect of small-molecule compounds, HEK293T cells were co-transfected with pEMX1/AAVS1/HPRT1-sgRNA, pD-EMX1/AAVS1/HPRT1, and HDR-USR vector. For SCR7 treatment, SCR7 (1 μM) was added 24 h after transfection until cells were collected. For nocodazole treatment, nocodazole (100 ng/mL) was added 12 h after transfection for 24 h, and released thereafter. Both groups were selected with puromycin for 5 days, beginning 48 h after transfection. Cells were harvested and pooled for subsequent isolation of gDNA for digestion assays.

Comparison of the Different Types of Donors on the Enrichment Efficiency of the HDR-USR System

For comparing conventional plasmid donors with ssODN donors, HEK293T cells were co-transfected with pEMX1-sgRNA, HDR-USR, and homologous donor templates, including conventional plasmid donor and ssODN donors of different lengths (89 nt, 110 nt, and 130 nt) (Figure 6D; Table S9), and selected with puromycin for 5 days. Cells were harvested, pooled, and used for isolating gDNA for digestion assays.

In addition, for comparing different types of double-stranded donors, HEK293T cells were co-transfected with pGOI-sgRNA; HDR-USR; and corresponding donors, including conventional plasmid donor pD-GOI, double cut plasmid donor pD-GOI-sg, linear donor containing a PCR-amplified sequence from pD-GOI, and double cut linear donor containing PCR product amplified from pD-GOI-sg. Subsequent cell collection and digestion assays were conducted as described earlier. GOI included *EMX1*, *AAVS1*, and *HPRT1*.

Statistics

Data obtained from the independent biological replicates of the experimental and the control groups were analyzed using an unpaired, two-tailed t test. p values less than 0.05 were considered statistically significant. All values were shown as mean ± SEM (standard error of the mean).

Availability of Data and Materials

The accession number for the deep sequencing raw data is BioProject: PRJNA516115.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.omtn.2019.12.021.

AUTHOR CONTRIBUTIONS

Z.Z., N.Y., and Y.S. designed the study. N.Y., Y.S., Y.F., J.D. and L.M. performed experiments and analyzed data. N.Y., Y.S., Z.Z., J.S.M., and

K.X. wrote the manuscript, with comments from all authors. All authors read and approved the final manuscript.

CONFLICTS OF INTEREST

The authors declare no competing interests.

ACKNOWLEDGMENTS

The authors thank all the colleagues in Z.Z.'s lab for their excellent technical assistance and helpful discussions. This study was supported by grants from the National Science and Technology Major Project of China (2018ZX08010-09B and 2014ZX08010-09B) and the National Natural Science Foundation of China (31702099).

REFERENCES

1. Maeder, M.L., and Gersbach, C.A. (2016). Genome-editing technologies for gene and cell therapy. Mol. Ther. *24*, 430–446.
2. Salsman, J., and Dellaire, G. (2017). Precision genome editing in the CRISPR era. Biochem. Cell Biol. *95*, 187–201.
3. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.
4. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823–826.
5. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.
6. Kanaar, R., Hoeijmakers, J.H., and van Gent, D.C. (1998). Molecular mechanisms of DNA double strand break repair. Trends Cell Biol. *8*, 483–489.
7. Pawelczak, K.S., Gavande, N.S., VanderVere-Carozza, P.S., and Turchi, J.J. (2018). Modulating DNA repair pathways to improve precision genome engineering. ACS Chem. Biol. *13*, 389–396.
8. Mao, Z., Bozzella, M., Seluanov, A., and Gorbunova, V. (2008). Comparison of nonhomologous end joining and homologous recombination in human cells. DNA Repair (Amst.) *7*, 1765–1771.
9. Knott, G.J., and Doudna, J.A. (2018). CRISPR-Cas guides the future of genetic engineering. Science *361*, 866–869.
10. Pinder, J., Salsman, J., and Dellaire, G. (2015). Nuclear domain 'knock-in' screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. Nucleic Acids Res. *43*, 9379–9392.
11. Shao, S., Ren, C., Liu, Z., Bai, Y., Chen, Z., Wei, Z., Wang, X., Zhang, Z., and Xu, K. (2017). Enhancing CRISPR/Cas9-mediated homology-directed repair in mammalian cells by expressing *Saccharomyces cerevisiae* Rad52. Int. J. Biochem. Cell Biol. *92*, 43–52.
12. Qi, Y., Zhang, Y., Zhang, F., Baller, J.A., Cleland, S.C., Ryu, Y., Starker, C.G., and Voytas, D.F. (2013). Increasing frequencies of site-specific mutagenesis and gene targeting in *Arabidopsis* by manipulating DNA repair pathways. Genome Res. *23*, 547–554.
13. Chu, V.T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., and Kühn, R. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat. Biotechnol. *33*, 543–548.
14. Canny, M.D., Moatti, N., Wan, L.C.K., Fradet-Turcotte, A., Krasner, D., Mateos-Gomez, P.A., Zimmermann, M., Orthwein, A., Juang, Y.C., Zhang, W., et al. (2018). Inhibition of 53BP1 favors homology-dependent DNA repair and increases CRISPR-Cas9 genome-editing efficiency. Nat. Biotechnol. *36*, 95–102.
15. Lin, S., Staahl, B.T., Alla, R.K., and Doudna, J.A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. eLife *3*, e04766.
16. Gutschner, T., Haemmerle, M., Genovese, G., Draetta, G.F., and Chin, L. (2016). Post-translational Regulation of Cas9 during G1 enhances homology-directed repair. Cell Rep. *14*, 1555–1566.
17. Baker, O., Tsurkan, S., Fu, J., Klink, B., Rump, A., Obst, M., Kranz, A., Schröck, E., Anastassiadis, K., and Stewart, A.F. (2017). The contribution of homology arms to nuclease-assisted genome engineering. Nucleic Acids Res. *45*, 8105–8115.
18. Ponce de León, V., Mérillat, A.M., Tesson, L., Anegón, I., and Hummler, E. (2014). Generation of TALEN-mediated GR^{dim} knock-in rats by homologous recombination. PLoS ONE *9*, e88146.
19. Böttcher, R., Hollmann, M., Merk, K., Nitschko, V., Obermaier, C., Philippou-Massier, J., Wieland, I., Gaul, U., and Förstemann, K. (2014). Efficient chromosomal gene modification with CRISPR/cas9 and PCR-based homologous recombination donors in cultured *Drosophila* cells. Nucleic Acids Res. *42*, e89.
20. Richardson, C.D., Ray, G.J., DeWitt, M.A., Curie, G.L., and Corn, J.E. (2016). Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat. Biotechnol. *34*, 339–344.
21. Yoshimi, K., Kunihiro, Y., Kaneko, T., Nagahora, H., Voigt, B., and Mashimo, T. (2016). ssODN-mediated knock-in with CRISPR-Cas for large genomic regions in zygotes. Nat. Commun. *7*, 10431.
22. Yang, L., Guell, M., Byrne, S., Yang, J.L., De Los Angeles, A., Mali, P., Aach, J., Kim-Kiselak, C., Briggs, A.W., Rios, X., et al. (2013). Optimization of scarless human stem cell genome editing. Nucleic Acids Res. *41*, 9049–9061.
23. Hisano, Y., Sakuma, T., Nakade, S., Ohga, R., Ota, S., Okamoto, H., Yamamoto, T., and Kawahara, A. (2015). Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish. Sci. Rep. *5*, 8841.
24. Zhang, J.P., Li, X.L., Li, G.H., Chen, W., Arakaki, C., Botimer, G.D., Baylink, D., Zhang, L., Wen, W., Fu, Y.W., et al. (2017). Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. Genome Biol. *18*, 35.
25. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308.
26. Li, K., Wang, G., Andersen, T., Zhou, P., and Pu, W.T. (2014). Optimization of genome engineering approaches with the CRISPR/Cas9 system. PLoS ONE *9*, e105779.
27. Steyer, B., Bu, Q., Cory, E., Jiang, K., Duong, S., Sinha, D., Steltzer, S., Gamm, D., Chang, Q., and Saha, K. (2018). Scarless genome editing of human pluripotent stem cells via transient puromycin selection. Stem Cell Reports *10*, 642–654.
28. Kim, H., Um, E., Cho, S.R., Jung, C., Kim, H., and Kim, J.S. (2011). Surrogate reporters for enrichment of cells with nuclease-induced mutations. Nat. Methods *8*, 941–943.
29. Ren, C., Xu, K., Liu, Z., Shen, J., Han, F., Chen, Z., and Zhang, Z. (2015). Dual-reporter surrogate systems for efficient enrichment of genetically modified cells. Cell. Mol. Life Sci. *72*, 2763–2772.
30. Niccheri, F., Pecori, R., and Conticello, S.G. (2017). An efficient method to enrich for knock-out and knock-in cellular clones using the CRISPR/Cas9 system. Cell. Mol. Life Sci. *74*, 3413–3423.
31. Ren, C., Xu, K., Segal, D.J., and Zhang, Z. (2019). Strategies for the enrichment and selection of genetically modified cells. Trends Biotechnol. *37*, 56–71.
32. Li, X., Bai, Y., Cheng, X., Kalds, P.G.T., Sun, B., Wu, Y., Lv, H., Xu, K., and Zhang, Z. (2018). Efficient SSA-mediated precise genome editing using CRISPR/Cas9. FEBS J. *285*, 3362–3375.
33. Wen, Y., Liao, G., Pritchard, T., Zhao, T.T., Connelly, J.P., Pruett-Miller, S.M., Blanc, V., Davidson, N.O., and Madison, B.B. (2017). A stable but reversible integrated surrogate reporter for assaying CRISPR/Cas9-stimulated homology-directed repair. J. Biol. Chem. *292*, 6148–6162.
34. Mitzelfelt, K.A., McDermott-Roe, C., Grzybowski, M.N., Marquez, M., Kuo, C.T., Riedel, M., Lai, S., Choi, M.J., Kolander, K.D., Helbling, D., et al. (2017). Efficient precision genome editing in iPSCs via genetic co-targeting with selection. Stem Cell Reports *8*, 491–499.
35. Shy, B.R., MacDougall, M.S., Clarke, R., and Merrill, B.J. (2016). Co-incident insertion enables high efficiency genome engineering in mouse embryonic stem cells. Nucleic Acids Res. *44*, 7997–8010.

36. Agudelo, D., Duringer, A., Bozoyan, L., Huard, C.C., Carter, S., Loehr, J., Synodinou, D., Drouin, M., Salsman, J., Dellaire, G., et al. (2017). Marker-free coselection for CRISPR-driven genome editing in human cells. Nat. Methods *14*, 615–620.

37. He, X., Tan, C., Wang, F., Wang, Y., Zhou, R., Cui, D., You, W., Zhao, H., Ren, J., and Feng, B. (2016). Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. Nucleic Acids Res. *44*, e85.

38. Sun, Y., Yan, N., Mu, L., Sun, B., Deng, J., Fang, Y., Shao, S., Yan, Q., Han, F., Zhang, Z., and Xu, K. (2019). sgRNA-shRNA structure mediated SNP site editing on porcine *IGF2* gene by CRISPR/StCas9. Front. Genet. *10*, 347.

39. Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature *533*, 420–424.

40. Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage. Nature *551*, 464–471.

41. Campla, C.K., Mast, H., Dong, L., Lei, J., Halford, S., Sekaran, S., and Swaroop, A. (2019). Targeted deletion of an NRL- and CRX-regulated alternative promoter specifically silences FERM and PDZ domain containing 1 (Frmpd1) in rod photoreceptors. Hum. Mol. Genet. *28*, 804–817.

42. Guo, T., Feng, Y.L., Xiao, J.J., Liu, Q., Sun, X.N., Xiang, J.F., Kong, N., Liu, S.C., Chen, G.Q., Wang, Y., et al. (2018). Harnessing accurate non-homologous end joining for efficient precise deletion in CRISPR/Cas9-mediated genome editing. Genome Biol. *19*, 170.

43. Mandal, P.K., Ferreira, L.M., Collins, R., Meissner, T.B., Boutwell, C.L., Friesen, M., Vrbanac, V., Garrison, B.S., Stortchevoi, A., Bryder, D., et al. (2014). Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. Cell Stem Cell *15*, 643–652.

44. Zheng, Q., Cai, X., Tan, M.H., Schaffert, S., Arnold, C.P., Gong, X., Chen, C.Z., and Huang, S. (2014). Precise gene deletion and replacement using the CRISPR/Cas9 system in human cells. Biotechniques *57*, 115–124.

45. Stafford, D.A., Dichmann, D.S., Chang, J.K., and Harland, R.M. (2017). Deletion of the sclerotome-enriched lncRNA PEAT augments ribosomal protein expression. Proc. Natl. Acad. Sci. USA *114*, 101–106.

46. Zhu, S., Li, W., Liu, J., Chen, C.H., Liao, Q., Xu, P., Xu, H., Xiao, T., Cao, Z., Peng, J., et al. (2016). Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. Nat. Biotechnol. *34*, 1279–1286.

47. Maruyama, T., Dougan, S.K., Truttmann, M.C., Bilate, A.M., Ingram, J.R., and Ploegh, H.L. (2015). Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat. Biotechnol. *33*, 538–542.

48. Hu, Z., Shi, Z., Guo, X., Jiang, B., Wang, G., Luo, D., Chen, Y., and Zhu, Y.S. (2018). Ligase IV inhibitor SCR7 enhances gene editing directed by CRISPR-Cas9 and ssODN in human cancer cells. Cell Biosci. *8*, 12.

49. Paix, A., Folkmann, A., Goldman, D.H., Kulaga, H., Grzelak, M.J., Rasoloson, D., Paidemarry, S., Green, R., Reed, R.R., and Seydoux, G. (2017). Precision genome editing using synthesis-dependent repair of Cas9-induced DNA breaks. Proc. Natl. Acad. Sci. USA *114*, E10745–E10754.

50. Yan, N., Xu, K., Li, X., Liu, Y., Bai, Y., Zhang, X., Han, B., Chen, Z., and Zhang, Z. (2015). Recombinant *Saccharomyces cerevisiae* serves as novel carrier for oral DNA vaccines in *Carassius auratus*. Fish Shellfish Immunol. *47*, 758–765.

51. Lee, J.S., Kallehauge, T.B., Pedersen, L.E., and Kildegaard, H.F. (2015). Site-specific integration in CHO cells mediated by CRISPR/Cas9 and homology-directed DNA repair pathway. Sci. Rep. *5*, 8572.