JBC RESEARCH ARTICLE

A Cas9–transcription factor fusion protein enhances homology-directed repair efficiency

Received for publication, December 10, 2020, and in revised form, March 3, 2021 Published, Papers in Press, March 6, 2021,
https://doi.org/10.1016/j.jbc.2021.100525

Guoling Li[‡], Haoqiang Wang[‡], Xianwei Zhang, Zhenfang Wu[*], and Huaqiang Yang[*]
From the College of Animal Science, South China Agricultural University, Guangzhou, China

Edited by Patrick Sung

Precise gene insertion or replacement in cells and animals that requires incorporation of a foreign DNA template into the genome target site by homology-directed repair (HDR) remains an inefficient process. One of the limiting factors for the inefficiency of HDR lies in the limited chance for colocalization of the donor template and target in the huge genome space. We here present a strategy to enhance HDR efficiency in animal cells by spatial and temporal colocalization of the donor and Cas9 by coupling the CRISPR system with a transcription factor (TF). We first identified that THAP domain-containing 11 (THAP11) can coordinate with CRISPR/Cas9 to increase HDR stably through screening multiple TFs from different species. We next designed donor structures with different fusion patterns with TF-specific DNA-binding motifs and found that appending two copies of THAP11-specific DNA binding motifs to both ends of the double-stranded donor DNA has an optimal effect to promote HDR. The THAP11-fused CRISPR system achieved more than twofold increase in HDR-mediated knock-in efficiency for enhanced green fluorescent protein (EGFP) tagging of endogenous genes in 293T cells. We also demonstrated up to 6-fold increases of knock-in through the combinational use of the TF-fused CRISPR and valnemulin, a recently discovered small-molecule HDR enhancer. This modified CRISPR system provides a simple but highly efficient platform to facilitate CRISPR-mediated KI manipulations.

The CRISPR/Cas9-induced precise genetic modification through homology-directed repair (HDR) is a desired process in many applications, but HDR is not readily performed for its low efficiency. The strategy enhancing HDR efficiency could benefit many HDR-involved applications, such as the creation of gene knock-in (KI) animals and gene therapy for correction of genetic mutations. HDR is a DNA recombination process that repairs double-strand break using a homologous donor as the repair template, thus requires a spatial and temporal colocalization of the template and target (1). It is anticipated that the donor searching for a homologous sequence is a key rate-limiting step for HDR occurrence because of the limited chance for coupling of the donor and target in the huge genome space. Therefore, increasing a local donor DNA concentration near the CRISPR/Cas9 cleavage site instead of allowing a randomly floating donor in the nucleoplasm could enhance the HDR frequency (2). This idea has been verified by studies using different strategies to physically position donor DNA near the cleavage site (2–8). In general, chemical modification of the donor DNA is required to mediate an attachment or binding between the donor and CRISPR components, such as using biotinylated donor and streptavidin-fused Cas9 protein to colocalize them (2, 6, 8), increasing the probability of the use of the donor as a repair template.

We here devised a simple system by using a nonchemically modified donor structure to colocalize donor and CRISPR components. The transcription factor (TF) controls gene transcription by binding to a specific DNA sequence, thereby can be exploited to recruit specific DNA in the presence of recognized DNA sequence (9). For this purpose, we constructed a fusion protein between Cas9 and TF DNA-binding domain (DBD) and appended the TF-recognized DNA sequence to the ends of donor to co-localize the two components in DNA repair process, with the aim to increase the HDR efficiency after CRISPR cleavage. This design is theoretically feasible, but some factors could affect its performance. For examples, is there sufficient affinity between the TF and DNA motif to realize a local enrichment of the donor DNA in the CRISPR cleavage site; how many and where are the binding motifs fused to the donor DNA; do the redesigned/fused protein and DNA structure affect the affinity and Cas9 cleavage activity? Therefore, identifying the TFs and binding motifs applicable in the CRISPR-mediated HDR system and optimizing their structures to realize an effective cooperation between them are critical.

In this study, we find that a TF, THAP domain-containing 11 (THAP11), can couple with the CRISPR system to effectively enhance HDR frequency (2–3 folds) when using a TF-binding motif-appended DNA donor. We also try a combinational use of the TF-fused CRISPR and small-molecule treatment and further increase the HDR frequency to 2- to 6-folds in dependence on the cell lines and genome sites tested.

‡ These authors contributed equally to this work.
* For correspondence: Zhenfang Wu, wzf@scau.edu.cn; Huaqiang Yang, yangh@scau.edu.cn.
Present address for Guoling Li: Institute of Neuroscience, State Key Laboratory of Neuroscience, Key Laboratory of Primate Neurobiology, CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai Research Center for Brain Science and Brain-Inspired Intelligence, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China.

Results

TF-fused Cas9 enables increased HDR efficiency

A total of 8 TFs, including human-, mouse-, and pig-derived activating transcription factor 3 (ATF3) and THAP11, as well as yeast GAL4 and PDR1, were tested with respect to their capacity to increase HDR in animal cells. We fused DBDs of these TFs (Table 1) to the C terminal of humanized spCas9. For the donor, we used a linear dsDNA with a 200-bp left homology arm and 800-bp right homology arm to insert a T2A–enhanced green fluorescent protein (EGFP) tag into the 3′ end of GAPDH gene, and two copies of TF-binding motifs were added to both ends of the dsDNA donor (Fig. 1*A*). The sequence information of TF-binding motifs was obtained from JASPAR 2020 (10) (Table 2). A predicted human ATF3 or THAP11-recognized sequence was used for all human, mouse, and pig TF orthologs. After cotransfection of Cas9–DBD/gRNA and donors harboring corresponding binding motifs, the HDR efficiency was determined by the EGFP tagging rate that successful tagging enables a fused GAPDH–EGFP expression. In all tested TFs, we found that only human- and mouse-derived THAP11 showed significant HDR-promoting effect (Fig. 1*B*). We further detected if different appending terminals of binding motif affect the HDR efficiency. Two copies of motifs were appended to either or both ends of dsDNA donors, and results demonstrated that a donor containing THAP11-binding motifs at both ends (LR) significantly and stably increased HDR efficiency when cotransfection with Cas9 fused with any of human, mouse, and pig THAP11–DBDs. However, the donor containing the same motifs in the right arm (R) only increased HDR in the presence of Cas9 fused with the DBD from mouse THAP11 (mTHAP11), and no effect was observed from the donor containing the motifs in left arm (L). Besides, coupled use of mouse ATF3 DBD-fused Cas9 and donor with ATF3 binding motifs at both ends also increased HDR, but the increase was not as prominent as that observed in THAP11 (Fig. 1*C*). We also performed Western blot to determine the EGFP levels in Cas9-mTHAP11-DBD transfection group and found the most prominent EGFP expression when using the donor with two end-appended motifs, followed by donors with motifs at the right end and then left end (Fig. 1*D*), in line with the HDR efficiency detected by flow cytometry in Figure 1*C*. All three modified versions of the donor increased EGFP expression level compared with the donor without binding motif (nonmodified control) (Fig. 1*D*). We next examined if increased numbers of the binding motif in the donor can increase HDR efficiency. To this purpose, we compared the donors with two and four copies of THAP11 motifs at both ends in the Cas9–mTHAP11–DBD system. Our findings showed that four copies of motifs slightly increased EGFP-positive cells compared with two copies of motifs in the donor, but the increase was insignificant (Fig. 1*E*). We also performed KI-specific PCR and southern blot, confirming that the EGFP-positive cells used to quantify HDR rate were all from HDR-mediated KI events without random integration of the donor into the genome (Fig. S1, *A*–*C*). These results collectively suggest that THAP11 is well applicable in the

Table 1
Gene sequences of DNA-binding domains of TFs used in this study

TFs	Sequences of DNA-binding domains	Species
ATF3	CAGGTCTCTGCCTCGGAAGTG	*Homo sapiens*
	AGTGCTTCTGCCATCGTCCCCTGC	
	CTGTCCCCTCCTGGGTCACTGGTGTTTG	
	AGGATTTTGCTAACCTGACGCCCTT	
	TGTCAAGGAAGAGCTGAGGTTTGC	
	CATCCAGAACAAGCACCTCTGCCACC	
	GGATGTCCTCTGCGCTGGAATCAGT	
	CACTGTCAGCGACAGACCCCTC	
	GGGGTGTCCATCACAAA	
	AGCCGAGGTAGCCCCTGAAGAAGATGA	
	AAGGAAAAAGAGGCGACG	
	AGAAAGAAATAAGATTGC	
	AGCTGCAAAG	
	CAACATCCAGGCCAGGTCTCTG	*Mus musculus*
	CCTCAGAAGTCAGTGCGACC	
	GCCATTGTCCCCTGCCTCTCAC	
	CTCCTGGGTCACTGGTAT	
	TTGAGGATTTTGCTAACCTGACA	
	CCCTTTGTCAAGGAAGAGCTG	
	AGATTCGCCATCCAGAATAAACA	
	CCTCTGCCATCGGATGTCCTCTGC	
	GCTGGAGTCAGTTACCGTCAACAA	
	CAGACCCCTGGAGATGTCAGTC	
	ACCAAGTCTGAGGCGGCCCCT	
	GAAGAAGATGAGAGGAAA	
	AGGAGGCGGCGAGAAAGAAAT	
	AAAATTGCTGCTGCCAAG	
	CAGGTCTCTGCCTCGGAAGTCAGTGC	*Sus scrofa*
	CTCTGCCATCGTCCCCTGCCTGTCC	
	CCTCCTGGGTCACTGGTGTTTGAAGA	
	TTTTGCTAACCTGACACCCTTTGTCA	
	AGGAAGAGCTGAGGTTCGCCATCCAG	
	AACAAGCACCTCTGCCACCGGATGTCCT	
	CGGCGCTGGAGTCGGTCACCGTCAG	
	CGGCAGACCCCTCGAGATGTCAGT	
	CACGAAAGCCGAGGTAGCCCCTGAAGAG	
	GATGAAAGGAAAAAGAGGAGA	
	CGGGAAAGAAATAAGATTGC	
	CGCCGCCAAG	
THAP11	ATGCCTGGCTTTACGTGCTGCGTG	*Homo sapiens*
	CCAGGCTGCTACAACAACTCG	
	CACCGGGACAAGGCGCTGCACT	
	TCTACACGTTTCCAAAGGACGCTG	
	AGTTGCGGCGCCTCTGGCTCAA	
	GAACGTGTCGCGTGCCGGCGTC	
	AGTGGGTGCTTCTCCACCTTCCA	
	GCCCACCACAGGCCACCGTCTC	
	TGCAGCGTTCACTTCCAGGGC	
	GGCCGCAAGACCTACACGGT	
	ACGCGTCCCCACCATCTTCCC	
	GCTGCGCGGCGTCAATGAG	
	CGCAAAGTAGCGCGCAGAC	
	CCGCTGGGGCCGCGGCCGCC	
	CGCCGCAGGCAGCAG	
	ATGCCTGGCTTTACGTGCTGCGT	*Mus musculus*
	TCCGGGCTGCTACAACAATTCA	
	CACCGGGACAAGGCGCTGCA	
	CTTCTACACGTTTCCCAAG	
	GACGCTGAGTTGCGGCGCCTC	
	TGGCTCAAGAACGTGTCCCGT	
	GCTGGCGTCAGTGGGTGCTT	
	CTCCACCTTCCAACCCACCACC	
	GGCCACCGTCTCTGCAGCGTCCA	
	CTTTCAGGGCGGCCGCAAGACC	
	TACACGGTGCGCGTTCCCACCA	
	TTTTCCCGCTGCGTGGCGTCAA	
	TGAGCGCAAAGTAGCTCGGAG	
	ACCTGCGGGAGCTGCGGCAG	
	CCCGCCGTAGGCAGCAG	
	ATGCCTGGCTTTACGTGCTGCG	*Sus scrofa*
	TGCCGGGCTGCTACAACAACT	
	CGCACCGGGATAAGGCGTTGCAC	
	TTCTACACGTTCCCCAAGGACGCTG	
	AGCTGCGGCGCCTCTGGCTCAA	
	GAATGTGTCCCGTGCTGGCGTC	
	AGTGGGTGCTTCTCCACCTTCC	
	AGCCCACCACAGGCCACCGTCT	
	CTGCAGCGTCCACTTCCAGGGCG	

Table 1—*Continued*

TFs	Sequences of DNA-binding domains	Species
	GCCGCAAAACCTACACCGTG CGGGTCCCCACCATCTTCCCG CTGCGCGGCGTCAACGAGCGC AAAGTAGCACGGCGACCCGC GGGGGCTGCAGCCGCCCGC CGCAGGCAGCAG	
	ATGCTGGCTTTACGTGCTGCGT TCCGGGCTGCTACAACAATTC ACACCGGGACAAGGCGCTGCAC TTCTACACGTTTCCCAAGGACGC TGAGTTGCGGCGCCTCTGGCTC AAGAACGTGTCCCGTGCTGGC GTCAGTGGGTGCTTCTCCACC TTCCAACCCACCACCGGCCAC CGTCTCTGCAGCGTCCACT TTCAGGGCGGCCGCAAGACC TACACGGTGCGCGTTCCCACCA TTTTCCCGCTGCGTGGCGTCA ATGAGCGCAAAGTAcGCTCGG AGACCTGCGGGAGCTGCGGCA GCCCGCCGTAGGCAGCAG	*Mus musculus* (Mutant inactivated mTHAP11)
GAL4	ATGAAGCTACTGTCTTCTATCGAA CAAGCATGCGATATTTGCCGA CTTAAAAAGCTCAAGTGCTCCA AAGAAAAACCGAAGTGCGCCAA GTGTCTGAAGAACAACTGGGAGT GTCGCTACTCTCCCAAAACCAA AAGGTCTCCGCTGACTAGGGCAC ATCTGACAGAAGTGGAATCAAGGC TAGAAAGACTGGAACAGCTATTTCTA CTGATTTTTCCTCGAGAAGACCT TGACATGATTTTGAAAATGGATT CTTTACAGGATATAAAAGCATTG TTAACAGGATTATTTGT ACAAGAT	*Saccharomyces cerevisiae*
PDR1	AAGCCAAGGTCGAAAGTAAG TAAAGCGTGCGATAACTGTAGA AAAAGAAAGATAAAATGTAATGG GAAGTTTCCCTGCGCAAGCTGTGA GATATATTCATGTGAGTGCACGTT CAGCACTAGACAAGGTGGC GCTCGAATA	*Saccharomyces cerevisiae*

TFs, transcription factors.

CRISPR/Cas9 system to enhance HDR, and the optimal effect can be achieved from coupled use of Cas9 fused with DBD from human THAP11 or mTHAP11 and the donor with two copies of motifs at both ends.

The above experiments established a reengineered Cas9 structure with HDR enhancement ability by fusion of Cas9 and THAP11–DBD. Accordingly, the donor should include two copies of THAP11-binding motifs at both ends to achieve an effective colocalization. The version of Cas9 fusion used in the following experiments is Cas9 fused with an mTHAP11-DBD-T2A-mCherry at the C terminal (Fig. 1*F*). mCherry can be used as the inner control to normalize the transfection efficiency to precisely quantify HDR efficiency. We also tested our system for EGFP tagging at the 3' end of hACTB gene by HDR in 293T cells and found a 2.2-fold increase in EGFP-positive cells and 2.6-fold increase in HDR efficiency by normalization of the transfection rate compared with using unfused Cas9 and the donor (Fig. 1*G*).

To test whether fusion with mTHAP11 affects the cleavage activity of Cas9, we used a single-strand annealing (SSA) assay to determine the targeted cleavage activity of the modified CRISPR system. The SSA reporter contains a mutant EGFP with a CRISPR cleavage site flanked by two tandem repeat sequences (Fig. 2*A*). Cleavage between the two repeats enables an SSA-mediated recombination to restore an intact EGFP sequence, therefore resuming the fluorescence that can be quantified by flow cytometry to report the CRISPR cleavage activity. We found that both Cas9 and Cas9-mTHAP11-DBD effectively cleaved the target site to generate the EGFP-positive cells with a similar rate (Fig. 2*B*). This result indicates that mTHAP11 fusion does not impair the targeted cleavage activity of Cas9.

TF fusion increases colocalization of the CRISPR and donor

To explore the mechanism of the TF-fused CRISPR in promoting HDR, we compared the expression levels of the key factors in DNA repair pathways between Cas9 and Cas9-mTHAP11-DBD edited 293T cells. Quantitative PCR results showed that the selected NHEJ factors, KU70 and PRKDC, and HDR factors, CtIP, RAD50, and RAD51, were not differentially expressed between Cas9 and Cas9-mTHAP11-DBD edited cells, implying that Cas9-mTHAP11-DBD does not promote HDR through affecting the DNA repair pathways (Fig. 3*A*). Next, to examine whether TF protects the donor DNA from degradation by their affinity to promote HDR, quantification of the donor DNA in transfected 293T cells showed no significant difference between the unfused and fused CRISPR system (Fig. 3*B*). The total cellular donor quantity is thus not affected in the presence of the TF-fused CRISPR.

We then replaced the WT mTHAP11–DBD with a mutant mTHAP11–DBD harboring a frame shift mutation in the C terminal. The mutant mTHAP11–DBD–fused CRISPR totally lost its ability in increasing HDR compared with the WT mTHAP11–DBD–fused CRISPR (Fig. 3*C*). To directly detect if TF fusion increases the binding between the CRISPR and donor, we used chromatin immunoprecipitation (ChIP) assay to evaluate the affinity between the TF-fused CRISPR and donor. The ChIP result showed a significantly increased binding efficiency between the CRISPR and donor after mTHAP11–DBD fusion compared with the unfused CRISPR system, whereas mutant mTHAP11-DBD had no increasing effect to their binding (Fig. 3*D*). In addition, we can observe an obvious colocalization of Cy3-labeled donor DNA and Cas9–mTHAP11–DBD protein in the nucleus of the transfected cells through immunofluorescence assay (Fig. 3*E*). These collective data indicate that the TF-fused CRISPR promotes HDR efficiency mainly through promoting colocalization of the donor and the CRISPR mediated by the affinity between the TF and its binding motif.

Characterization of novel robust small-molecule HDR enhancers

Small-molecule treatment is a simple approach to increase HDR in animal cells (11–14). We here also tried to use small molecules to further increase HDR efficiency. Previously reported small-molecule enhancers usually have a marginal or nonidentical effect to promote HDR (15, 16). Some commonly used small-molecule HDR enhancers are also cytotoxic agents, thereby exerting a negative effect in application in cells or animals. To identify new small molecules with the

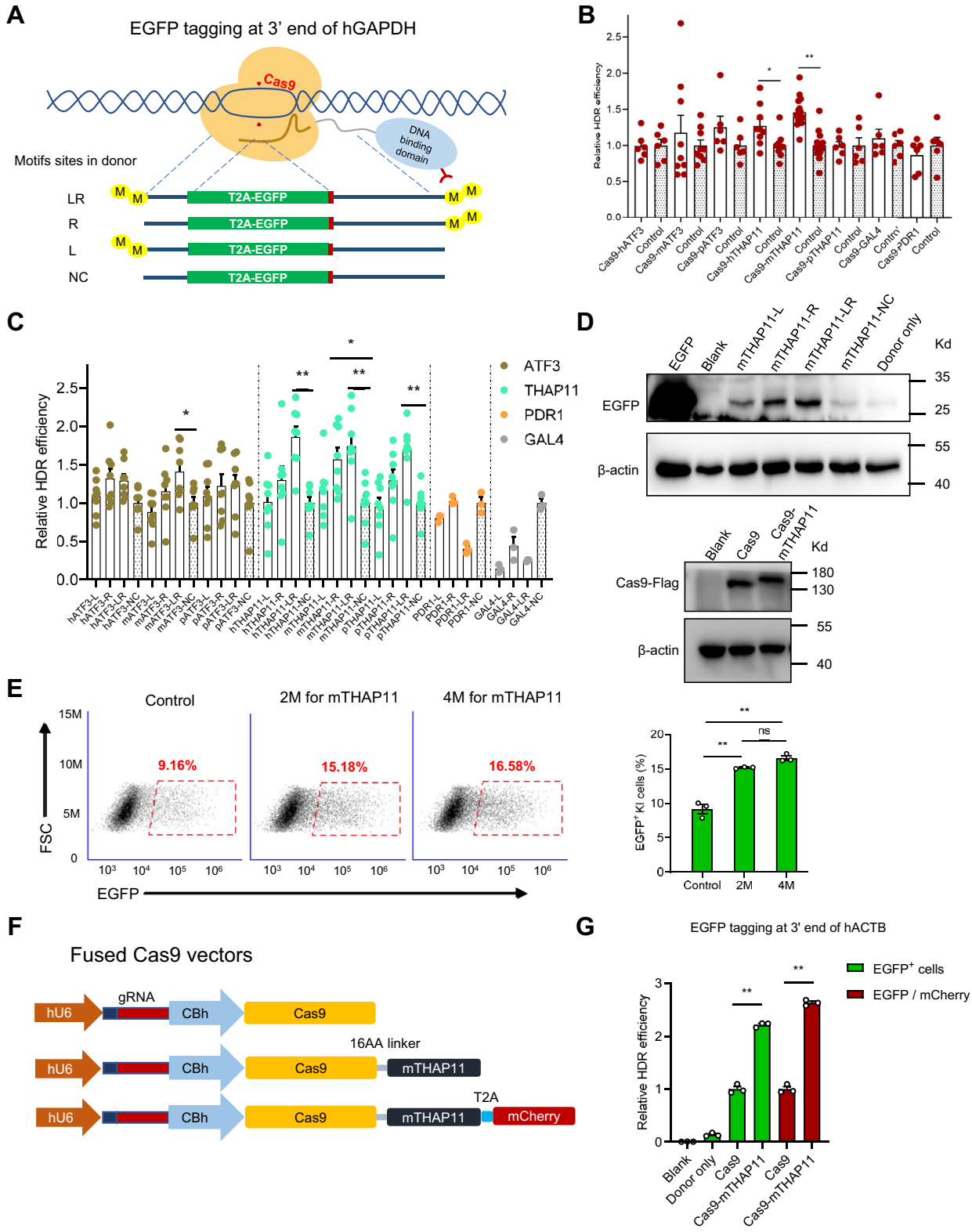

Figure 1. Coupling CRISPR and TF increases HDR efficiency in 293T cells. *A*, strategy of using TF-fused Cas9 for EGFP tagging at the end of hGAPDH gene in 293T cells. *B*, testing HDR-promoting effect of different TFs. The 293T cells were transfected with Cas9, which was fused with eight TFs, including human, mouse, and pig-derived ATF3 and THAP11 as well as yeast GAL4 and PDR1, gRNA recognizing the hGAPDH targeting site, and the donor appending with two copies of corresponding TF-binding motifs. EGFP-positive cells were determined by flow cytometry and normalized to mCherry-denoted transfection efficiency. Values are expressed as fold changes relative to controls. Controls are cells transfected with unmodified CRISPR and donor. *C*, the effect of the tagging position of TF-binding motifs in the donor on HDR efficiency. Two copies of TF-binding motifs were appended to either left (L) or right (R) end, or both ends (LR), and transfected with the corresponding TF-fused CRISPR to analyze the HDR efficiency in 293T cells. HDR is expressed as the percentage of EGFP-positive cells normalized to the percentage of mCherry-positive cells. Values are shown as fold changes relative to NC donor + corresponding TF-fused CRISPR. *D*, determination of tagged EGFP by Western blot in the mTHAP11-fused CRISPR system. *Upper panel*, comparison of donors

Table 2
Binding site motifs of TFs used in this study

TFs	Binding site motifs[a]	Species[b]
THAP11	AGGACTACATTTCCCAGCA	*Homo sapiens*
ATF3	GGTGACGTCATC	*Homo sapiens*
GAL4	CGGGGAACAGTACTC	*Saccharomyces cerevisiae*
PDR1	TCCGCGGA	*Saccharomyces cerevisiae*

TFs, transcription factors.
[a] Predicted TF binding motifs according to JASPAR 2020 (http://jaspar.genereg.net) (10).
[b] The species origin of TF used for prediction of binding motifs.

HDR-promoting effect but less toxicity, our previous work screened a U.S. Food and Drug Administration-approved drug library and found a series of new HDR enhancers (unpublished data). Among them, two clinically available drugs (U.S. Food and Drug Administration or European Medicines Agency approval), vinblastine and valnemulin, showing robust HDR-promoting effects, were further tested in the present study. Vinblastine, as a mitotic inhibitor, is a clinically important antitumor agent, and valnemulin is an antibiotic for veterinary use (17). In a dose-dependent assay for testing the HDR-promoting effect, vinblastine showed its maximal effect at 5 μM, while valnemulin had the similar promoting effect at the dosage range of 1 to 10 μM (Fig. 4*A*). Combinational use of 5-μM vinblastine and 1-μM valnemulin can additively increase HDR compared with either of them alone (Fig. 4*B*). Testing the cytotoxicity of the two drugs at their optimized concentrations found that valnemulin did not affect cell viability and apoptosis levels but vinblastine severely decreased cell viability and increased cell apoptosis in 293T cells, compared with dimethyl sulfoxide-treated controls (Fig. 4, *C* and *D*). The possible effect of valnemulin on CRISPR cleavage was tested by the T7 Endonuclease I (T7E1) assay (Fig. S2*A*). After CRISPR transfection and valnemulin treatment in 293T cells, no significant difference was found in the CRISPR cleavage activity, shown by a similar T7E1 cleavage efficiency (Fig. S2*A*) and similar proportion of mutant alleles (Fig. S2*B*) between valnemulin treatment and control groups.

Great HDR improvement by the combinational use of TF-fused Cas9 and small molecule

To examine the combinational effect of Cas9–mTHAP11 and small molecule on increasing HDR, 293T cells were cotransfected with the CRISPR vector containing Cas9–mTHAP11 and gRNAs and donors for EGFP tagging at GAPDH/ACTB, and then treated with 1-μM valnemulin for 48 h. Analyzing HDR efficiency through flow cytometry found that the use of Cas9–mTHAP11 increased HDR efficiency 4 and 2 times greater at GAPDH and ACTB loci, respectively, compared with Cas9 transfection (Fig. 5, *A* and *B*). Furthermore, valnemulin treatment had additional 0.4- and 1.6-fold increased efficiency for the two loci (Fig. 5, *A* and *B*), indicating an additional benefit for enhancement of HDR in the TF-fused CRISPR system. We also tested the combination of Cas9–mTHAP11 and valnemulin in other cell lines and found great increase in HDR efficiency. By testing in the EGFP tagging system at GAPDH and ACTB loci, approximately 2- and 1.6-fold increased HDR efficiency was found in HeLa cells (Fig. 5*C*), and approximately 4-fold increase for both loci was achieved in HepG2 cells (Fig. 5*D*). These HDR-promoting effects were greatly higher than using either of Cas9–mTHAP11 or valnemulin alone.

Discussion

Increasing local concentration of donor DNA at the CRISPR/Cas9 cleavage site is a feasible way to promote CRISPR/Cas9–mediated HDR efficiency. Some recent works have achieved significantly increased HDR in animal cells and embryos *via* multiple similar approaches (2–8). For examples, using an avidin-fused Cas9 to recruit biotin-modified single-stranded oligodeoxynucleotide (ssODN) donor (500- to 800-nt fragments flanked by 100-nt homology arms) showed 15 to 22% KI efficiency in mouse embryos, whereas no KI editing was observed in all tested genes for the unmodified CRISPR and donor system (2). Similar avidin–biotin conjugation strategy achieved 43 to 95% KI efficiency for large-fragment dsDNA KI in two cell–stage mouse embryos, compared with 11.2 to 43.2% KI efficiency when using the unmodified CRISPR system (6). Ling *et al.* developed a chemically modified Cas9 protein containing azide-containing noncanonical amino acid, which allows conjugation of dibenzylcyclooctyne (DBCO)-containing donor DNA by means of strain-promoted alkyne-azide cycloaddition, thereby recruiting the donor template to the cleavage complex. Cotransfection of noncanonical amino acid–modified Cas9–gRNA ribonucleoprotein and DBCO-modified ssODN had a 1.6-fold HDR increase as compared with unmodified groups in 293 cells. When using a DBCO-modified DNA adaptor to tether unmodified donor ssODN by base pairing, modified ribonucleoprotein–adaptor

harboring two copies of mTHAP11-binding motifs in either or both ends showing that donor with binding motifs at both ends (mTHAP11-LR) had the highest EGFP tagging rate, followed by using the donor with binding motifs at right end (mTHAP11-R) and then left end (mTHAP11-L). *Lower panel*, Western blot showing successful expression of mTHAP11-fused Cas9 (Cas9–mTHAP11) in 293T cells. mTHAP11-fused Cas9 has a higher molecular weight than unmodified Cas9 protein. EGFP denotes an EGFP-expressing plasmid transfection; blank, nontransfected cells. No CRISPR was used in the donor-only group. *E*, the effect of the number of TF-binding motifs in the donor on HDR efficiency. Donors with two and four copies of mTHAP11-binding motifs (shown as 2M and 4M) at both ends were compared as respect to their ability to enhance HDR, in cotransfection with Cas9–mTHAP11. A similar HDR efficiency represented by the EGFP-positive rate ($EGFP^+$ KI rate) can be found between donors with two and four binding motifs. NC donor was used in control. *F*, structure of the optimized mTHAP11-fused Cas9 vector. A U6-driven gRNA expression cassette and a chicken β-actin hybrid (CBh) promoter-driven Cas9 were assembled in the same vector for coexpression. Reengineered Cas9 were fused with mTHAP11 DBD and then tagged with a mCherry in the C-terminal to monitor transfection efficiency. *G*, the mTHAP11-fused CRISPR system enhances the EGFP tagging rate at the 3′ end of hACTB gene in 293T cells. Flow cytometry assay showing increased EGFP-positive cells and normalized HDR efficiency in mTHAP11-fused CRISPR-transfected cells compared with unmodified CRISPR. Values are expressed as fold changes relative to the unfused Cas9 group. Blank, nontransfected cells. No CRISPR was used in the donor-only group. $*p < 0.05$ and $**p < 0.01$ compared with controls (unmodified Cas9 and donors). ATF3, activating transcription factor 3; DBD, DNA-binding domain; HDR, homology-directed repair; L, left homology arm; M, TF-binding motifs; mTHAP11, mouse THAP11; NC, nonmodified control donor; ns, not significant; R, right homology arm; TF, transcription factor.

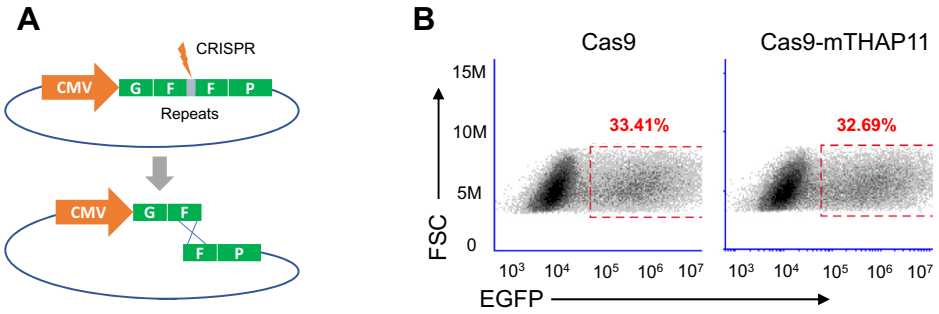

Figure 2. Impact of TF fusion on the Cas9 cleavage activity. *A*, the schematic diagram of the SSA assay detecting Cas9 cleavage efficiency. *B*, Cas9–mTHAP11 and unmodified Cas9 showed similar targeted cleavage activity represented by SSA-mediated EGFP repair efficiency. mTHAP11, mouse THAP11; SSA, single-strand annealing; TF, transcription factor.

conjugates increased HDR efficiency approximately 5-fold as compared with unmodified or scramble adapters (7). These studies confirm that methods copositioning the CRISPR cleavage complex and donor template can induce a notable increase in HDR efficiency.

Based on the similar strategy, we colocalized the CRISPR and donor by utilizing the affinity between the TF and its binding motif. We screened multiple pairs of TFs and binding motifs and found that THAP11 of different species origin in fusion with Cas9 protein all increased CRISPR-mediated HDR.

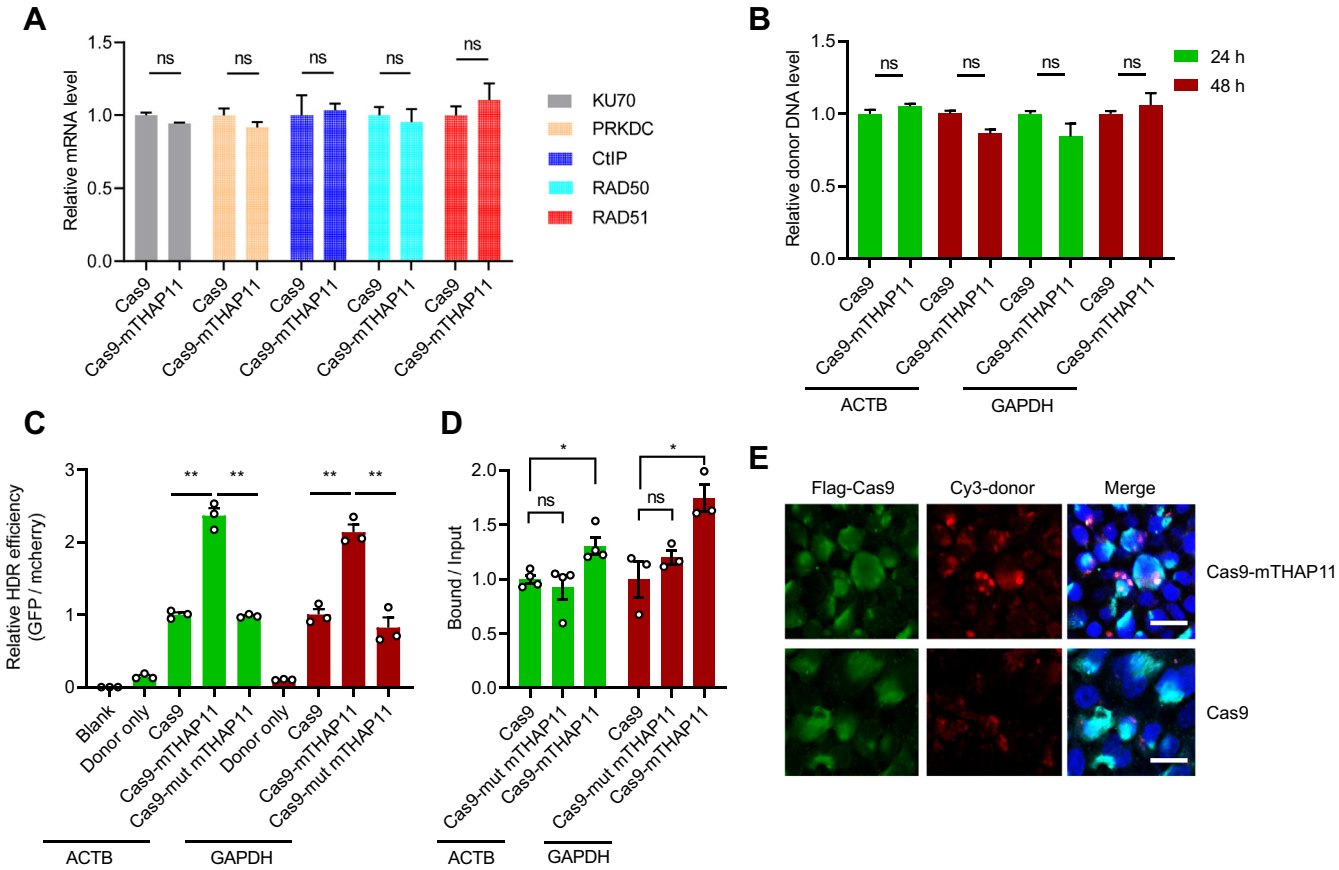

Figure 3. TF fusion promotes HDR through increasing colocalization of the CRISPR and donor. *A*, the expression level of key factors of NHEJ pathways (KU70 and PRKDC) and HDR factors (CtIP, RAD50, and RAD51) does not differ significantly with either TF-fused Cas9 or unfused Cas9 transfection. Values are fold changes to unfused Cas9 groups. *B*, quantification of the total cellular donor DNA level showing a similar quantity of donor at 24 and 48 h after transfection, indicating that TF fusion does not affect donor concentration within cells. Values are fold changes to unfused Cas9 groups. *C*, mutant mTHAP11 has no HDR-promoting effect when fused with the CRISPR. HDR is expressed as the percentage of EGFP-positive cells normalized to the percentage of mCherry-positive cells. Values are shown as fold changes relative to the corresponding unfused Cas9 group. *D*, the ChIP assay showing increased binding ability between the TF-fused CRISPR and donor. However, mutation in the TF reversed the increasing effect on their binding. *E*, colocalization of Cy3-labeled donor DNA (*red fluorescence*) and TF-fused Cas9 protein (*green fluorescence*) can be observed in the nuclei (*blue fluorescence*) of transfected 293T cells but rarely found in unfused Cas9-transfected cells. *$p < 0.05$ and **$p < 0.01$ between two groups. The scale bars represent 10 μm. ChIP, chromatin immunoprecipitation; HDR, homology-directed repair; mTHAP11, mouse THAP11; ns, not significant; TF, transcription factor.

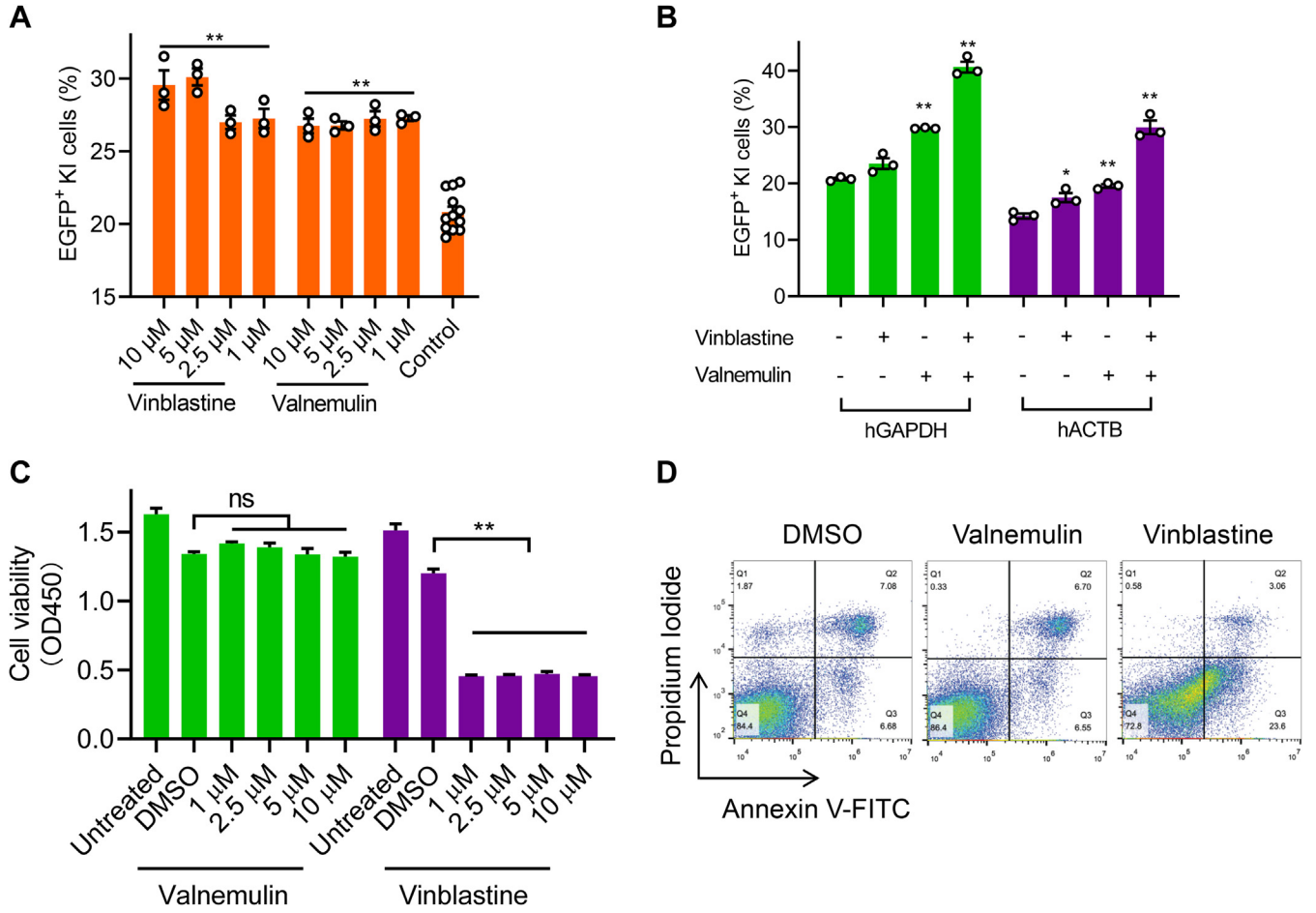

Figure 4. Valnemulin and vinblastine are effective HDR enhancers. *A*, valnemulin and vinblastine at the dose range of 1 to 10 μM demonstrated a significantly enhanced HDR rate in 293T cells using the unmodified CRISPR system and donor for EGFP KI in the GAPDH locus shown in Figure 1*A*. *B*, combinational treatment of the two small molecules showed a further enhanced HDR rate compared with using either of them alone. *C*, cell viability tested by the CCK-8 assay after 48-h treatment of small molecules (1- to 10-μM valnemulin and 1- to 10-μM vinblastine) in 293T cells. *D*, cell apoptosis tested by annexin V and propidium iodide double staining after 24-h treatment of small molecules (10-μM valnemulin and 10-μM vinblastine) in 293T cells. The control shown in panel *A* is DMSO-treated 293T cells, and untreated groups shown in panel *C* are 293T without any small molecules and DMSO treatment. *$p < 0.05$ and **$p < 0.01$ compared with DMSO control. CCK-8, Cell Counting Kit-8; DMSO, dimethyl sulfoxide; HDR, homology-directed repair; KI, knock-in; ns, not significant.

The best promoting effect can be achieved in cotransfection of THAP11 DBD-fused Cas9 and dsDNA donor tagged with two tandem copies of the binding motif at both ends. However, the donor tagged with just one-end binding motif did not significantly increase HDR efficiency, implying that increased numbers of binding motifs at the donor increase affinity between the donor and TF-fused Cas9. Among the other tested TFs, mATF3 also had tendency to increase HDR, but yeast-derived TFs, PDR1 and GAL4, showed no effect on HDR enhancement, indicating an incompatibility in TF usage between yeast and mammals. Compared with previous reports that generally used chemically modified Cas9 and donor DNA, our genetically modified method greatly reduced workload and cost in preparing the CRISPR system and donors. The modified CRISPR vector can be directly used in transfection, and the modified donor can be conveniently assembled by molecular cloning. Moreover, our system works on the dsDNA donor, which allows HDR with a long template, thereby expanding the applicability for various KI manipulations.

Another finding of this study is that the TF-fused CRISPR/donor system can be compatible with small molecules to further increase HDR efficiency. The new identified small molecules were clinically available medicines, thus enabling a safe application in animal models or clinics. In particular, valnemulin is an antiinfection agent that inhibits protein translation in bacteria by binding peptidyl transferase enzyme in 50s ribosomal subunit (17). No report shows its function in eukaryotic cells. We also observed that valnemulin has no cytotoxicity in 293T cells, shown by unaffected cell viability and apoptosis levels compared with dimethyl sulfoxide control. These results indicate the potential safety of using valnemulin in animal models and clinical investigation for KI enhancement purpose. By combinational use of TF-fused CRISPR/Cas9 system and valnemulin, approximately 6-fold increase in the KI rate can be found in 293T cells, compared with only the use of unmodified CRISPSR/Cas9 system. The combinational use also increased KI rates to approximately 2 and 4 times in HeLa and HepG2 cells, respectively, indicating a broad

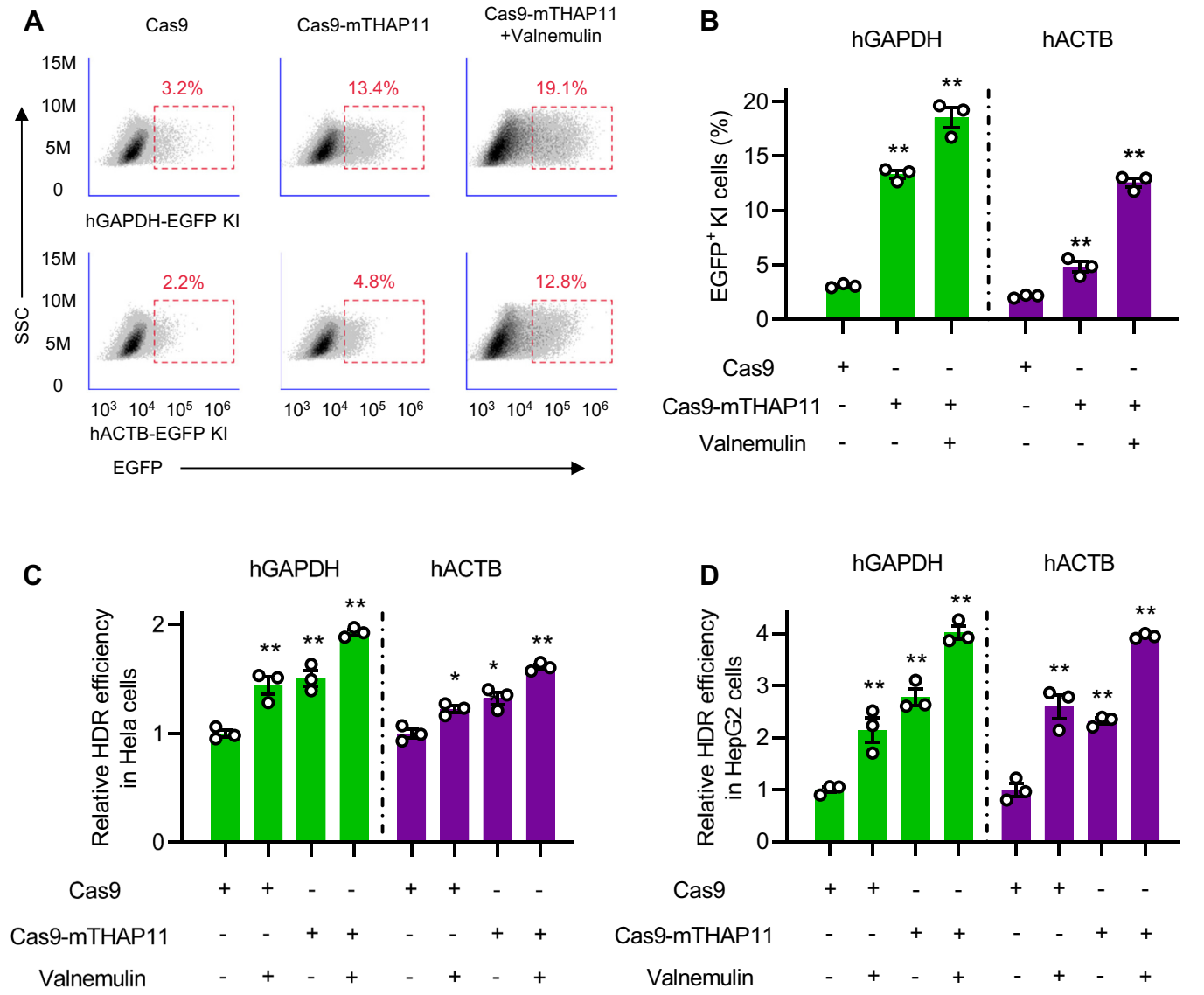

Figure 5. Combination of the mTHAP11-fused CRISPR and small molecule greatly increases HDR in multiple cells. *A*, representative flow cytometry results for CRISPR-mediated EGFP tagging at GAPDH and ACTB loci in 293T showing Cas9–mTHAP11 increased HDR and the combinational use of Cas9–mTHAP11 and 1-μM valnemulin showed additive benefit to further increase HDR. *B*, summarized results of panel *A* showing approximately 6-fold increase in the HDR rate at GAPDH and ACTB loci by Cas9–mTHAP11 and valnemulin combination, compared with using unmodified Cas9 only. *C* and *D*, testing the combinational effect of Cas9–mTHAP11 and valnemulin in HeLa and HepG2 cells. Cas9–mTHAP11 and valnemulin combination showed the highest HDR rates for CRISPR-mediated EGFP tagging at GAPDH and ACTB loci in the 2 cell lines. *$p < 0.05$ and **$p < 0.01$ compared with control cells (unmodified Cas9 transfection and no drug treatment). HDR, homology-directed repair; mTHAP11, mouse THAP11.

HDR-promoting effect of our system in various cell lines. In summary, we developed a TF-fused CRISPR/Cas9 system as a universal platform that facilitates an enhanced KI manipulation in animal cells. This system is easy-to-construct, applicable in multiple cell lines, and can work in synergy with small-molecule treatment to markedly promote KI efficiency.

Experimental procedures

Construction of the TF-modified CRISPR/Cas9 system

The TF-modified Cas9 gene was constructed by fusion of DBDs of selected TFs (ATF3, THAP11, GAL4, and PDR1) to the 3′ terminal of humanized SpCas9, with a 16AA linker (SGSETPGTSESATPES) between them. To avoid influence of inconsistent transfection efficiency on quantifying HDR, we also included a mCherry that tags to Cas9-TF-DBD by a T2A sequence. The genome editing efficiency can be calibrated by normalizing to the mCherry-positive rate. Donor templates are dsDNA containing an inserted T2A–EGFP sequence flanked by homology arms, and two or four copies of specific TF-binding motifs at the ends of homology arms. Cotransfection of the donor and TF-modified Cas9 into animal cells can mediate a precise insertion of T2A–EGFP to the end of specific gene (GAPDH or ACTB), forming an EGFP-tagged fusion gene. All vectors were assembled by the In-Fusion HD cloning system (Takara).

Cell culture, transfection, and drug treatment

The 293T cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum at 37 °C

in air containing 5% CO_2. All cell culture reagents were purchased from Thermo Fisher Scientific. For KI manipulation, 293T cells were seeded in 24-well plates and transfected with Lipofectamine 3000 reagent (Thermo Fisher Scientific), according to the manufacturer's manual. For drug treatment, small molecules (vinblastine sulfate and valnemulin hydrochloride, both are products of MCE) were added into the cell culture at 24 h after transfection, and cells were further incubated for 48 h for analysis. The concentrations of vinblastine and valnemulin are 5 and 1 μM, respectively, for cell treatment, except a gradient concentration (1, 2.5, 5, and 10 μM) shown in Figure 4.

Quantitation of KI efficiency

KI or HDR efficiency of EGFP tagging at specific genes was detected by analyzing the EGFP-positive cells by flow cytometry (BD Accuri C6, BD Biosciences) after 72-h transfection or 48-h drug treatment as mentioned above. In case of using vectors including mCherry fusion Cas9, mCherry-positive cells were also gated to present transfection efficiency. KI efficiency was calculated by normalizing the EGFP rate to mCherry rate.

Western blot

Western blot was performed as reported previously (18). In brief, 293T cells were lysed using RIPA buffer (Thermo Fisher Scientific), and the lysate supernatant was collected by centrifugation. The supernatant with an equal amount of the total protein was loaded on SDS-PAGE for protein separation. Afterward, separated protein in SDS-PAGE gel was transferred onto the PVDF membrane. The target proteins were labeled with primary antibodies of Flag (20543-1-AP, Proteintech), EGFP (50430-2-AP, Proteintech), and β-actin (20536-1-AP, Proteintech), followed by reaction with HRP-conjugated secondary antibodies. The protein signal was developed using Pierce ECL substrate (Thermo Fisher Scientific) and photographed using an UVP imaging system.

KI genotyping

Precise KI of the T2A–EGFP sequence into ACTB or GAPDH loci was verified by PCR and southern blot. For PCR identifying KI, the primer pairs covering the homology arm but located in the genomic region outside the homology arm were used (Fig. S2*A*). Southern blot was performed according to a previous report (19). Genomic DNA from KI 293T cells was digested with PstI, which flanks to the inserted sites of T2A–EGFP to generate 2733-bp and 1950-bp KI sequences for ACTB and GAPDH loci, respectively. Any other untargeted integration would generate positive bands with different lengths. The probe was a digoxigenin-labeled PCR product through amplification of the partial EGFP region.

Quantitative PCR

Quantitative PCR was performed to assay the mRNA expression levels of selected NHEJ and HDR key factors. Total RNA of transfected cells (Cas9/Cas9-mTHAP11 and donor cotransfection) were extracted using the TRIzol reagent (Thermo Fisher Scientific) and converted into cDNA as PCR templates through reverse transcription using the PrimeScript RT reagent Kit with gDNA Eraser (Takara). Quantitative PCR was performed using PowerUp SYBR Green Master Mix (Thermo Fisher Scientific) and QuantStudio 6 Flex Real-Time PCR System (Thermo Fisher Scientific). The relative mRNA abundance of each gene was calculated by the $2^{-\Delta\Delta Ct}$ method, normalized to endogenous β-actin expression level, and expressed as the fold change as compared with the unfused Cas9 groups.

Immunofluorescence

Donor DNA targeting to the ACTB locus was labeled with Cy3 and cotransfected into 293T cells with the CRISPR vector harboring Cas9 or Cas9–mTHAP11 and U6-expressing ACTB gRNA. At 24 h after transfection, cells were fixed with 4% paraformaldehyde, permeated with 0.3% Triton X-100, blocked in 5% bovine serum albumin, and stained with anti-Flag antibody (20543-1-AP, Proteintech) to label Cas9 protein. The cells were then incubated with the Alexa fluor 488–conjugated secondary antibody (Thermo Fisher Scientific) and counterstained with 4',6-diamino-2-phenylindole. The fluorescent signal was examined under an inverted fluorescence microscope to observe the cellular location of Cas9 and the donor.

ChIP

We used ChIP to evaluate the binding ability between Cas9-mTHAP11-DBD protein and mTHAP11-binding-motif-appeneded DNA donor. ChIP was performed according the manufacturer's manual (ChIP-IT High Sensitivity Kit, Active motif). In brief, 293T cells were transfected with mTHAP11-binding-motif-appeneded DNA donor and CRISPR plasmids containing Cas9-mTHAP11-DBD and gRNA targeting GAPDH and ACTB, respectively. The cells were fixed and then lysed to prepare chromatin at 72 h after transfection. Chromatin preparations were sheared by sonication, and Cas9/Cas9-mTHAP11-DBD protein was immunoprecipitated by Flag antibody (SAB4301135, Sigma-Aldrich) and Protein G agarose beads. Afterward, the ChIP DNA was purified and used as the template for quantitative PCR. The sheared DNA before immunoprecipitation was used as the input. The following primer pairs were used for quantification of the ACTB donor and GAPDH donor, respectively: hACTB-donor-F: CTCCATCGTCCACCGCAAA, hACTB-donor-R: GAACAGCTCCTCGCCCTTG, and hGAPDH-donor-F: ACAACGAATTTGGCTACAGCAAC, hGAPDH-donor-R: GAACAGCTCCTCGCCCTTG. They were also used to amplify ACTB or GAPDH donors of input DNA for normalization purpose.

CRISPR cleavage assay

CRISPR cleavage efficiency was determined by the SSA-based assay or T7E1 digestion (20). The SSA assay is an HDR after formation of a double-strand break (21). We introduced a CRISPR-cleavage site flanked by two repeated sequence in the same direction into EGFP driven by a human cytomegalovirus immediate early enhancer and promoter. After CRISPR cleavage, SSA-mediated repair of the two

repeats restored an intact EGFP coding region. As SSA repair is a highly efficient DNA repair process, CRISPR-targeted cleavage efficiency can be determined by quantifying EGFP-positive cells with flow cytometry. For the T7E1 assay, the repaired region was amplified by PCR after CRISPR-mediated gene editing. The PCR products were denatured and annealed to form mismatches, which can be cleaved by T7E1 (New England Biolabs), as nonhomologous end-joining repair of Cas9-induced breaks will leave a variety of different mutations at the target site. T7E1 was then added into the annealed DNA solution for digestion of mismatches. CRISPR cleavage efficiency was represented by the ratio of mutant PCR products which can be digested by T7E1 in the total PCR product.

Cell viability and apoptosis assay

Changes in cell viability after small molecule treatment were measured by using the Cell Counting Kit-8 (Dojindo). Tested cells were seeded in 96-well plates and treated with vinblastine and valnemulin at the indicated dosages for 48 h. Cell culture was then replaced with 100 μl of the fresh medium, and then 10 μl of the Cell Counting Kit-8 solution was directly added to the culture medium to detect cell viability. Plates were incubated for approximately 2 h at 37 °C, and absorbance was measured at 450 nm using a microplate reader (Tecan Spark). Cell apoptosis after small-molecule treatment was analyzed by double staining of Annexin V-FITC and propidium iodide (Keygen), which can identify apoptotic cells and dead cells, respectively. After staining, cells were measured with NovoCyte flow cytometer (ACEA), and data were analyzed with FlowJo X software.

Statistical analysis

All data are presented as the mean ± SEM and calculated from values of independent experiments, which are labeled as circles or dots on each bar in the plots unless otherwise indicated in figure legends. Statistical significance was determined by Student's t test analysis (two-tailed) for two groups and one-way ANOVA with Dunnett's post hoc test for three or more groups (GraphPad Prism 8). Differences were considered statistically significant at $p < 0.05$.

Data availability

All data are contained within the article.

Supporting information—This article contains supporting information.

Author contributions—G. L., H. Y., and Z. W. conceived the project. G. L., H. Y., H. W., and X. Z. designed and performed the experiments. H. Y. and G. L. drafted and edited the manuscript. All authors reviewed the results and approved the final version of the manuscript.

Funding and additional information—This work was supported in part by the National Natural Science Foundation of China (31772555).

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: ATF3, activating transcription factor 3; ChIP, chromatin immunoprecipitation; DBCO, dibenzylcyclooctyne; DBD, DNA-binding domain; EGFP, enhanced green fluorescent protein; HDR, homology-directed repair; KI, knock-in; mTHAP11, mouse THAP11; SSA, single-strand annealing; ssODN, single-stranded oligodeoxynucleotide; T7E1, T7 Endonuclease I; TF, transcription factor; THAP11, THAP domain-containing 11.

References

1. Jasin, M., and Rothstein, R. (2013) Repair of strand breaks by homologous recombination. *Cold Spring Harb. Perspect. Biol.* **5**, a012740
2. Ma, M., Zhuang, F., Hu, X., Wang, B., Wen, X. Z., Ji, J. F., and Xi, J. J. (2017) Efficient generation of mice carrying homozygous double-floxp alleles using the Cas9-avidin/biotin-donor DNA system. *Cell Res.* **27**, 578–581
3. Lee, K., Mackley, V. A., Rao, A., Chong, A. T., Dewitt, M. A., Corn, J. E., and Murthy, N. (2017) Synthetically modified guide RNA and donor DNA are a versatile platform for CRISPR-Cas9 engineering. *Elife* **6**, e25312
4. Savic, N., Ringnalda, F. C., Lindsay, H., Berk, C., Bargsten, K., Li, Y., Neri, D., Robinson, M. D., Ciaudo, C., Hall, J., Jinek, M., and Schwank, G. (2018) Covalent linkage of the DNA repair template to the CRISPR-Cas9 nuclease enhances homology-directed repair. *Elife* **7**, e33761
5. Aird, E. J., Lovendahl, K. N., St Martin, A., Harris, R. S., and Gordon, W. R. (2018) Increasing Cas9-mediated homology-directed repair efficiency through covalent tethering of DNA repair template. *Commun. Biol.* **1**, 54
6. Gu, B., Posfai, E., and Rossant, J. (2018) Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. *Nat. Biotechnol.* **36**, 632–637
7. Ling, X., Xie, B., Gao, X., Chang, L., Zheng, W., Chen, H., Huang, Y., Tan, L., Li, M., and Liu, T. (2020) Improving the efficiency of precise genome editing with site-specific Cas9-oligonucleotide conjugates. *Sci. Adv.* **6**, eaaz0051
8. Carlson-Stevermer, J., Abdeen, A. A., Kohlenberg, L., Goedland, M., Molugu, K., Lou, M., and Saha, K. (2017) Assembly of CRISPR ribonucleoproteins with biotinylated oligonucleotides via an RNA aptamer for precise gene editing. *Nat. Commun.* **8**, 1711
9. Spitz, F., and Furlong, E. E. (2012) Transcription factors: From enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613–626
10. Fornes, O., Castro-Mondragon, J. A., Khan, A., van der Lee, R., Zhang, X., Richmond, P. A., Modi, B. P., Correard, S., Gheorghe, M., Baranašić, D., Santana-Garcia, W., Tan, G., Chèneby, J., Ballester, B., Parcy, F., *et al.* (2020) JASPAR 2020: Update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **48**, D87–D92
11. Song, J., Yang, D., Xu, J., Zhu, T., Chen, Y. E., and Zhang, J. (2016) RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat. Commun.* **7**, 10548
12. 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
13. 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
14. 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
15. Riesenberg, S., and Maricic, T. (2018) Targeting repair pathways with small molecules increases precise genome editing in pluripotent stem cells. *Nat. Commun.* **9**, 2164

16. 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., Xu, J., Chun, N., Yuan, W., Cheng, T., and Zhang, X. B. (2017) Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol.* **18**, 35
17. Stipkovits, L., Ripley, P. H., Tenk, M., Glávits, R., Molnár, T., and Fodor, L. (2005) The efficacy of valnemulin (Econor) in the control of disease caused by experimental infection of calves with Mycoplasma bovis. *Res. Vet. Sci.* **78**, 207–215
18. Yang, H., Wang, G., Sun, H., Shu, R., Liu, T., Wang, C. E., Liu, Z., Zhao, Y., Zhao, B., Ouyang, Z., Yang, D., Huang, J., Zhou, Y., Li, S., Jiang, X., *et al.* (2014) Species-dependent neuropathology in transgenic SOD1 pigs. *Cell Res.* **24**, 464–481
19. Zhang, X., Li, Z., Yang, H., Liu, D., Cai, G., Li, G., Mo, J., Wang, D., Zhong, C., Wang, H., Sun, Y., Shi, J., Zheng, E., Meng, F., Zhang, M., *et al.* (2018) Novel transgenic pigs with enhanced growth and reduced environmental impact. *Elife* **7**, e34286
20. Li, H. L., Gee, P., Ishida, K., and Hotta, A. (2016) Efficient genomic correction methods in human iPS cells using CRISPR-Cas9 system. *Methods* **101**, 27–35
21. Pâques, F., and Haber, J. E. (1999) Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. *Microbiol. Mol. Biol. Rev.* **63**, 349–404