Cas9-Rep fusion tethers donor DNA *in vivo* and boosts the efficiency of HDR-mediated genome editing

Zhentao Zhou[1,2,†], Jiahui Xiao[1,2,†], Shuai Yin[1,2,†], Yache Chen[1,2], Yang Yuan[1], Jianwei Zhang[1] (iD), Lizhong Xiong[1] (iD) and Kabin Xie[1,2,*] (iD)

[1]*National Key Laboratory of Crop Genetic Improvement, Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan, China*

[2]*Hubei Key Laboratory of Plant Pathology, Huazhong Agricultural University, Wuhan, China*

Summary

Genome editing based on the homology-directed repair (HDR) pathway enables scar-free and precise genetic manipulations. However, the low frequency of HDR hinders its application in plant genome editing. In this study, we engineered the fusion of Cas9 and a viral replication protein (Rep) as a molecular bridge to tether donor DNA *in vivo*, which enhances the efficiency of targeted gene insertion via the HDR pathway. This Rep-bridged knock-in (RBKI) method combines the advantages of rolling cycle replication of viral replicons and *in vivo* enrichment of donor DNA at the target site for HDR. Chromatin immunoprecipitation indicated that the Cas9-Rep fusion protein bound up to 66-fold more donor DNA than Cas9 did. We exemplified the RBKI method by inserting small- to middle-sized tags (33–519 bp) into 3 rice genes. Compared to Cas9, Cas9-Rep fusion increased the KI frequencies by 4–7.6-fold, and up to 72.2% of stable rice transformants carried in-frame knock-in events in the T_0 generation. Whole-genome sequencing of 6 plants segregated from heterozygous KI lines indicated that the knock-in events were faithfully inherited by the progenies with neither off-target editing nor random insertions of the donor DNA fragment. Further analysis suggested that the RBKI method reduced the number of byproducts from nonhomologous end joining; however, HDR-mediated knock-in tended to accompany microhomology-mediated end joining events. Together, these findings show that the *in vivo* tethering of donor DNAs with Cas9-Rep is an effective strategy to increase the frequency of HDR-mediated genome editing.

Introduction

The clustered regularly interspaced palindromic repeat (CRISPR)/CRISPR-associated protein 9 (Cas9) system has been engineered as a versatile platform for genome editing (Gao, 2021; Wang and Doudna, 2023; Zhou *et al*., 2023). Cas9 and guide RNA (gRNA) are programmed to cut a genomic target and edit gene sequences via the endogenous DNA repair machinery. Double-stranded DNA breaks (DSBs) are predominantly repaired by nonhomologous DNA end joining (NHEJ), which is prone to introducing insertions and deletions (InDels) for targeted gene mutagenesis. Homology-directed repair (HDR) requires donor DNA as the template to repair DSBs via homologous recombination. HDR can be engineered to precisely insert and replace the DNA sequence at the target site (Miki *et al*., 2018) but has an extremely low frequency (Yeh *et al*., 2019). In some cases, DSBs are repaired by microhomology-mediated end-joining (MMEJ), and this pathway has been used to generate large deletions for plant genome engineering (Tan *et al*., 2020). Additionally, the prime editor (PE) engineers the reverse transcriptase to insert short sequences using the pegRNA as a template (Anzalone *et al*., 2019). The integration of site-specific recombinase (Sun *et al*., 2024) and transposase (Liu *et al*., 2024) into the CRISPR/Cas9 system can result in the insertion of a large DNA fragment. Among different DSB repairing pathways, HDR enables precise and scar-free gene insertion and

replacement; however, the low efficiency of HDR hinders its utilization in genome editing (Chen *et al*., 2022; Dong and Ronald, 2021).

Many efforts have been made to increase the frequency of HDR-mediated genome editing in plants. Biolistic bombardment (Dong *et al*., 2020; Sun *et al*., 2016) and viral vectors (Baltes *et al*., 2014; Cermak *et al*., 2015; Endo *et al*., 2016; Gil-Humanes *et al*., 2017; Wang *et al*., 2017) have been used to increase the abundance of donor DNAs in recipient cells to increase HDR frequencies. In biolistic bombardment, the stability of donor dsDNA or single-stranded oligo DNA nucleotide (ssODN) templates is further enhanced by the addition of phosphorothioate linkages (Lu *et al*., 2020). For viral vectors, wheat dwarf virus (WDV) and bean yellow dwarf virus (BeYDV) have been engineered for HDR-mediated gene targeting (Baltes *et al*., 2014; Cermak *et al*., 2015; Endo *et al*., 2016; Gil-Humanes *et al*., 2017; Timmermans *et al*., 1992; Wang *et al*., 2017). These methods increase the overall copy number of free donor DNA to increase the availability of DNA templates for HDR. In addition to donor DNA concentrations in recipient cells, DNA-template mobility and proximity are also important for HDR (Renkawitz *et al*., 2014). Various strategies have been developed in animal systems to increase the HDR template at the DSB site by tethering donor DNA and the Cas9 protein *in vitro*. Several protein tags have been used for this purpose, including avidin

(Carlson-Stevermer *et al.*, 2017; Ma *et al.*, 2017), SNAP (Savic *et al.*, 2018), HUH from porcine circovirus 2 (PCV) replication protein (Rep) (Aird *et al.*, 2018; Lovendahl *et al.*, 2017) and the recombinant Cas9 protein containing a genetically encoded noncanonical amino acid (Ling *et al.*, 2020). In plants, the fusion of Cas9 and the *Agrobacterium* VirD2 protein was utilized to tether an HDR template *in vitro* for biolistic bombardment or *Agrobacterium*-mediated transformation, resulting in increased KI frequencies (Ali *et al.*, 2020; Tang *et al.*, 2023). Recently, Nagy *et al.* (2022) reported that particle bombardment of *in vitro*-conjugated Cas9-HUH and 70 nt-ssODN complexes could increase the KI frequency in soybean. These approaches are restricted by the tedious procedures of DNA–protein bioconjugation *in vitro*, the limitation of ssODN length and the inefficient delivery of protein–DNA complexes via biolistic bombardment transformation. To date, HDR-mediated scar-free KI is still a major challenge in plant genome editing.

We envision that increasing the spatial–temporal availability of donor DNA could boost the frequency of HDR. Notably, the viral Rep protein nicks the geminiviral genome and remains covalently linked to the viral strand DNA to initiate rolling cycle replication *in vivo* (Figure S1) (Bonnamy *et al.*, 2023; Chandler *et al.*, 2013; Hanley-Bowdoin *et al.*, 2013; Laufs *et al.*, 1995; Tompkins *et al.*, 2021). Thus, we hypothesized that the Rep protein could be engineered to tether DNA molecules *in vivo* to increase HDR frequencies for precise KI in genome editing. In this study, we hijack the WDV Rep protein to enrich donor DNA for the Cas9/gRNA complex (Figure 1), achieving scar-free KI frequencies of up to 72.2%. Given its simplicity and versatility, we anticipate

that this method will be broadly used in plant gene targeting and potentially enable other types of HDR-mediated genetic manipulations.

Results

Design of a Rep-based bridge for *in vivo* tethering of donor DNA and Cas9

The geminiviral Rep protein specifically nicks the replication origin in the large intergenic region (LIR) sequence and is covalently linked to the viral ssDNA during rolling cycle replication (see Figure S1 for schematics of WDV replication) (Bonnamy *et al.*, 2023; Chandler *et al.*, 2013; Hanley-Bowdoin *et al.*, 2013; Laufs *et al.*, 1995; Tompkins *et al.*, 2021). This inspired us to engineer the Rep protein as a molecular bridge that can carry donor DNA to the genome editing machinery. We hypothesized that the fusion of the Rep and Cas9 proteins could tether the DNA template *in vivo* and then cleave the chromosomal target, which would increase the probability of HDR (Figure 1). This method is referred to hereafter as Rep-bridged knock-in (RBKI).

To implement RBKI, we incorporated the WDV replicon and its Rep protein into a CRISPR/Cas9 genome editing vector (Xie *et al.*, 2015). The resulting vectors contain a WDV replicon to carry the HDR template, a Cas9-Rep expression cassette, and a gRNA expression cassette (Figure 2a). The WDV cassette contains a *Rep* gene, two LIR fragments, a small intergenic region (SIR), and a target-specific HDR template that replaces the viral genome fragment encoding the capsid protein and movement protein
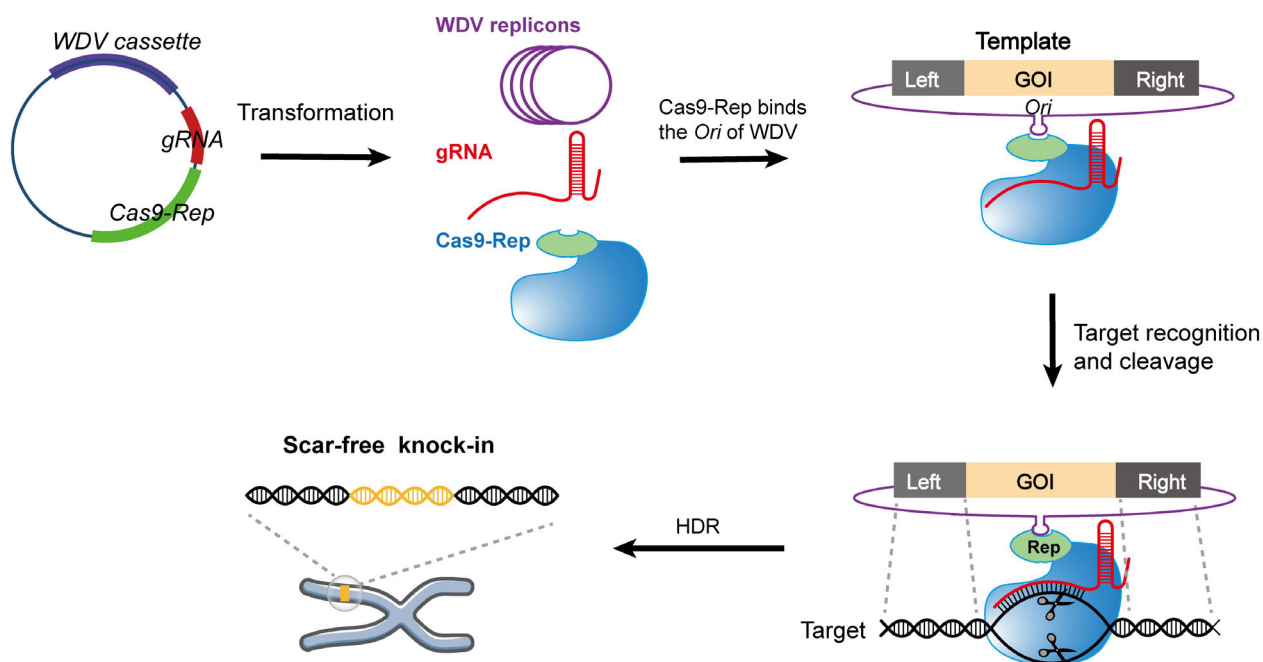
Figure 1 Design of a Rep-based protein bridge tethering the donor DNA and Cas9 *in vivo*. The viral replicon (e.g. WDV), gRNA and Cas9-Rep cassettes are included in a plasmid vector and delivered into plant cells (e.g. Agrobacterium-mediated transformation). The geminiviral replicon (e.g. WDV) carries the donor template in which the gene of interest (GOI) is flanked by target-specific left and right homologous arms for HDR. In transformed cells, Cas9-Rep and gRNA are expressed while the WDV is amplified by rolling-cycle replication. The fusion of Cas9 and the Rep protein nicks and covalently links the stem–loop sequence at the replication origin (*Ori*). The ternary complex of Cas9-Rep, the donor template, and gRNA, binds to the chromosomal target and cleaves the target DNA to trigger the repair of DSBs. The donor template linked to the Cas9-Rep protein serves as the template for HDR, resulting in precise and scar-free KI with high efficiency.
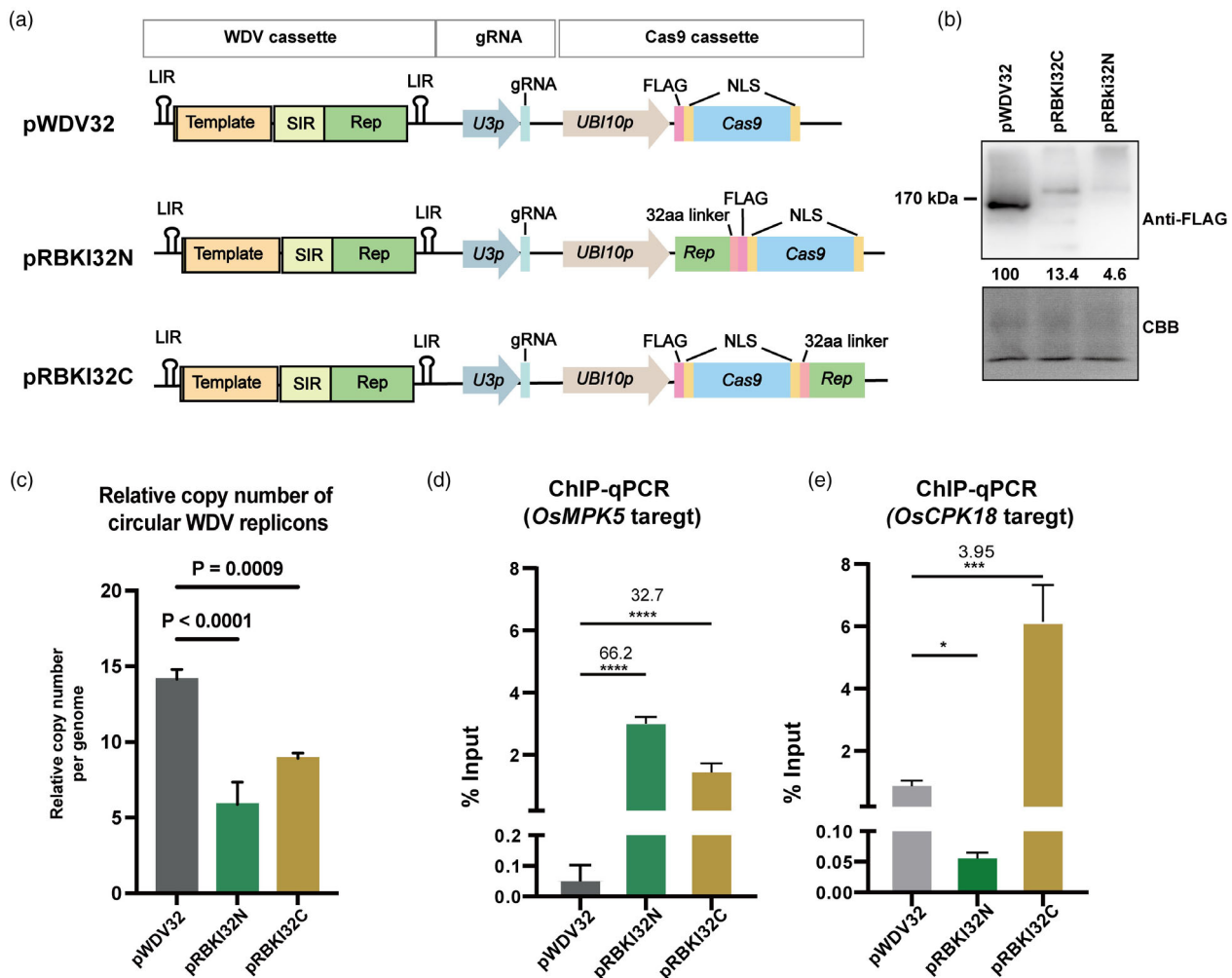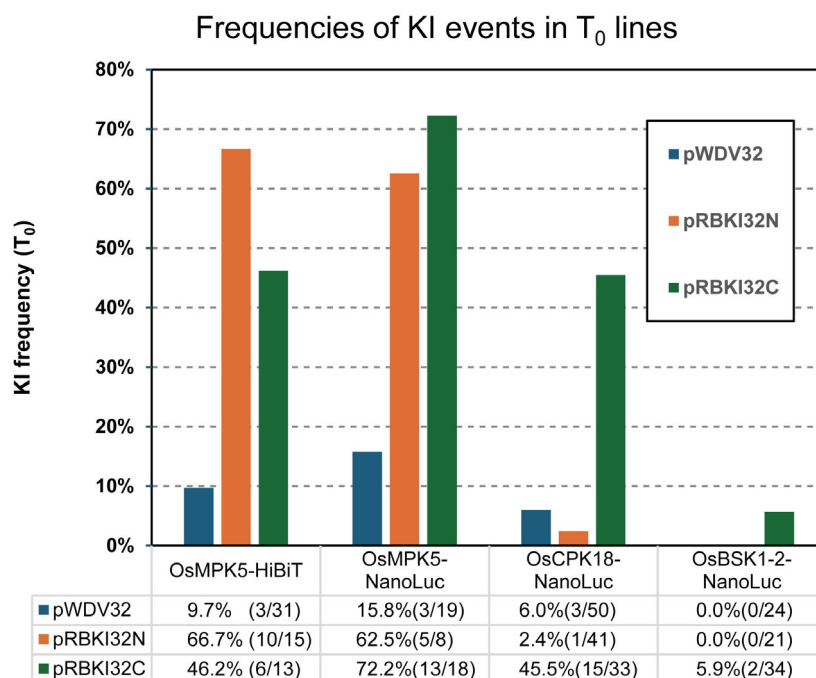
Figure 2 Efficient enrichment of the donor DNA template with Cas9-Rep. (a) Schematic of the vector structure for RBKI vectors. *U3p*, snoRNA U3 promoter; *UBI10p*, rice *UBIQUITIN 10* promoter; *NLS*, nuclear localization signal. (b) Western blot showing the expression of the Cas9 (pWDV32), Cas9-Rep (pRBKI32C) and Rep-Cas9 (pRBKI32N) proteins in rice protoplasts. Anti-FLAG antibodies were used to detect the Cas9 and fusion proteins. The numbers below the lanes indicate the relative protein levels estimated from the band intensities. CBB, Coomassie Brilliant Blue stained total protein. (c) Relative copy number of circular WDV DNA in protoplasts. The copy numbers of circular WDV replicons relative to the rice genome are shown. Error bar, standard deviation (n = 3 technical replicates). (d,e) ChIP–qPCR analysis of interactions between WDV DNA and Cas9-Rep fusion proteins at the *OsMPK5* and *OsCPK18* targets. The chromatin DNA was immunoprecipitated with Cas9, Cas9-Rep, and Rep-Cas9, followed by qPCR with WDV-specific primers. The ChIP signals were calculated as the IP/input ratio (% Input) for each sample. The number on top of the bars represents the relative enrichment of WDV replicons using pWDV32 samples as references. ****, *$P < 0.0001$, 0.05 (Student's t test, n = 3 technical replicates).

(Gil-Humanes *et al*., 2017; Matzeit *et al*., 1991). Because both the N- and C-termini of Cas9 are amenable to fusing effector domains, we generated pRBKI32N and pRBKI32C vectors, which fused the full-length Rep gene to the N- and C-termini of Cas9 (referred to as Rep-Cas9 and Cas9-Rep), respectively (Figure 2a). We also generated a pWDV32 vector that contains the original Cas9 cassette for side–by–side comparisons with the RBKI strategy. After the pRBKI constructs are delivered into plant cells (e.g. *Agrobacterium* transformation), the cyclized WDV replicons were initially produced from the T-DNAs and then amplified by rolling cycle replication. We hypothesized that the Cas9-Rep fusion protein captures the free extrachromosomal WDV replicons and carries the DNA template in WDV to the chromosomal target. Cas9/gRNA-mediated cleavage of target DNA triggers HDR using the proximal donor template, resulting in scar-free gene insertion or replacement (Figure 1). We expect that

this method could increase the spatial–temporal availability of donor DNA to boost the efficiency of HDR genome editing.

Cas9-Rep efficiently binds WDVs carrying the HDR template in rice protoplasts

We first examined the mRNA transcription of the fused genes because the WDV *Rep* gene has an 86 bp intron and translates into two proteins, namely, the full-length Rep and the truncated RepA proteins from intron-splicing and intron-retaining transcripts, respectively (Schalk *et al*., 1989). Rep and RepA form oligomers and are required for the initiation of roll cycle replication (Missich *et al*., 2000). In rice protoplasts, reverse transcription–PCR (RT–PCR) identified an additional intron (135 bp) of *Rep* that has not been previously reported (Figure S2). Although *Rep* has 4 splicing forms in rice, the desired Cas9- and Rep-fused proteins were detected for the pRBKI32N

Frequencies of KI events in T_0 lines

	OsMPK5-HiBiT	OsMPK5-NanoLuc	OsCPK18-NanoLuc	OsBSK1-2-NanoLuc
■ pWDV32	9.7% (3/31)	15.8%(3/19)	6.0%(3/50)	0.0%(0/24)
■ pRBKI32N	66.7% (10/15)	62.5%(5/8)	2.4%(1/41)	0.0%(0/21)
■ pRBKI32C	46.2% (6/13)	72.2%(13/18)	45.5%(15/33)	5.9%(2/34)

Figure 3 KI efficiencies of the RBKI method at 3 targets in stable rice transformants (T_0 generation). The RBKI vectors (pRBKI32N and pRBKI32C) and pWDV32 were used to insert the HiBiT tag or NanoLuc into *OsMPK5*, *OsCPK18*, and *OsBSK1-2*. The primary transformants that carried KI events were compared (see all genotyping results in Figures S6–S9). The numbers in brackets indicate the number of KI-positive lines and total transformants (KI/Total).

and pRBKI32C vectors (Figure 2b). However, compared with those of the Cas9 protein expressed from pWDV32, the Cas9-Rep and Rep-Cas9 proteins from the two RBKI vectors were reduced by 86.6% and 95.4%, respectively, in protoplast transient expression assays (Figure 2b), likely due to the alternative splicing of the *Rep* gene (Figure S2). Though the fusion of Cas9 and Rep had lower expression, NHEJ-introduced InDels were detected in stable rice transformants when these two RBKI vectors were used to target a rice gene (Figure S3), suggesting that both fusion proteins retained target cleavage activity for genome editing.

The RBKI vectors simultaneously express Rep and Cas9-Rep fusion proteins, and the Rep proteins function as oligomers (Missich *et al.*, 2000); therefore, Cas9-Rep fusion proteins may interfere with WDV replication. To test this possibility, the relative copy numbers of the WDV replicons were compared between the pRBKI and pWDV32 vectors via quantitative PCR (qPCR) (Figure S4). In rice protoplasts, both the pRBKI and pWDV32 vectors generated circularly closed DNA replicons, as expected (Figure S4). Compared with pWDV32, the two RBKI vectors yielded approximately 2 times fewer circular WDV replicons (Figure 2c), indicating that the Cas9-Rep fusion protein slightly impaired the replication of WDV.

To validate the feasibility of *in vivo* tethering of DNAs via the RBKI approach, the rice *MITOGEN-ACTIVATED PROTEIN KINASE 5* (*OsMPK5*) and *CALCIUM-DEPENDENT PROTEIN KINASE 18* (*OsCPK18*) were targeted via pRBKI32N/C and pWDV32 vectors (see target sequences and gRNA design in Figure S5). We used chromatin immunoprecipitation coupled with quantitative real-time PCR (ChIP–qPCR) to quantify the WDV replicons tethered to the Cas9-Rep proteins in rice protoplasts. The ChIP signals were quantified using the percentage input method (Solomon *et al.*, 2021). For the *OsMPK5* target, the ChIP signals of

Rep-Cas9 (pRBKI32N) and Cas9-Rep (pRBKI32C) were 66.2- and 32.7-fold higher than that of Cas9 (pWDV32) (Figure 2d, Student's *t* test, $P < 1.0 \times 10^{-4}$). For the *OsCPK18* target, Cas9-Rep efficiently enriched WDV DNA with 3.95-fold higher ChIP signals than Cas9 (pWDV32), but the Rep-Cas9 fusion failed to enrich donor DNA (Figure 2e). The inconsistent enrichment of WDV replicons with Rep-Cas9 (*OsMPK5* target vs *OsCPK18* target) was likely due to the different donor template sequences that were loaded in the WDV cassette. Though the Cas9-rep protein and WDV replicons were reduced (Figure 2b,c), the ChIP–qPCR results indicated that Cas9-Rep stably enriched WDV replicons, as we hypothesized.

The RBKI method efficiently increases HDR-mediated KI frequencies

To examine whether the RBKI strategy increases HDR frequency, pWDV32, pRBKI32N and pRBKI32C were used for in-frame insertion of a 33 bp HiBiT tag and 519 bp nanoluciferase (NanoLuc) before the stop codon of *OsMPK5* (Figure S5). NanoLuc has extremely high luminescence and enables in-gel detection of tagged proteins (Hall *et al.*, 2012; Li *et al.*, 2021). HiBiT is a small fragment from NanoLuc that can be detected in membranes without antibodies (Schwinn *et al.*, 2018). The KI events of *OsMPK5::HiBiT* and *OsMPK5::NanoLuc* (double colon denotes insertion hereafter) were examined via insertion-specific PCR and Sanger sequencing (Figures S6 and S7). Although *NanoLuc* is 16 times longer than *HiBiT,* their KI frequencies did not significantly differ. The KI frequencies of *HiBiT* were 66.7% and 46.2%, whereas the KI frequencies of *NanoLuc* were 62.5% and 72.2% for pRBKI32N and pRBKI32C, respectively (Figure 3). In contrast, for the KI frequencies of the pWDV32 vector, only 9.7% and 15.8% of the T_0 lines carried KI events for *HiBiT* and
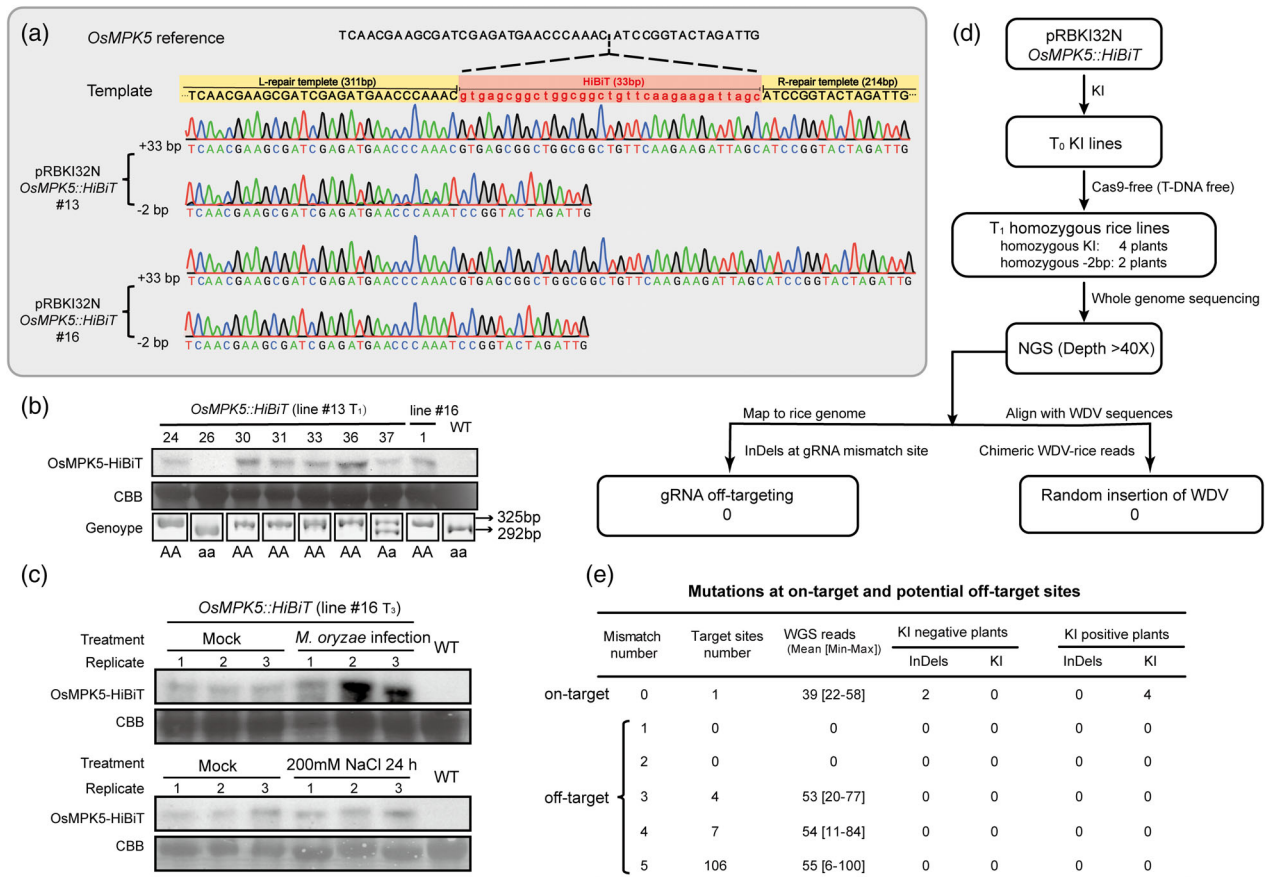
Figure 4 Analysis of the inheritance of KI events and examination of off-target editing for *OsMPK5::HiBiT*. (a) Genotypes of *OsMPK5::HiBiT* lines #13 and #16. Sanger sequencing revealed that both plants carried monoallelic HiBiT-KI (+33 bp) and a 2 bp deletion (−2 bp) at the *OsMPK5* target. (b) HiBiT blotting analysis of *OsMPK5::HiBiT* T_1 plants. AA, homozygous KI; Aa, heterozygous KI; aa, homozygous −2 bp deletion. (c) HiBiT blotting analysis of *OsMPK5-HiBiT* expression after *M. oryzae* infection and high-salinity treatment. Positive KI rice plants at T_3 generations were used for the treatments. CBB, Coomassie Brilliant Blue stained total protein. WT, wild-type plants (cv. KitaaKe). (d, e) Off-target editing was analysed in 2 KI-negative and 4 KI-positive plants (T-DNA free) via whole-genome sequencing (WGS). (d) Flowchart showing the procedure of WGS analysis. (e) InDels and KIs detected via WGS. Genomic sites paired with guide sequences with 1–5 mismatches were considered potential off-target sites of gRNAs. The numbers in the last 4 columns (InDel and KI) indicate the number of editing events observed in 6 plants on the basis of the WGS data.

NanoLuc (Figure 3), respectively. For both tags, the RBKI method increased the HDR-mediated KI frequencies by 4–6-fold at the *OsMPK5* target.

We further used the RBKI method to knock in *NanoLuc* before the stop codon of rice *OsCPK18* and *BRASSINOSTEROID-SIGNALLING KINASE 1–2* (*OsBSK1-2*) (Figure S5). For the *OsCPK18* target, the pRBKI32C vectors resulted in 7.6 times higher KI frequencies than did pWDV32 (45.5% vs. 6.0%, Figures 3 and S8). However, the pRBKI32N vector was less efficient, which is consistent with the ChIP–qPCR results that Rep-Cas9 did not enrich WDV replicons for the *OsCPK18* target (Figure 2e). For the *OsBSK1-2* target, the pRBKI32C vector, with a Cas9-Rep configuration, resulted in a KI frequency of 5.9%, whereas no KI lines were obtained via pWDV32 (Figures 3 and S9). pRBKI32N showed comparable KI frequencies to those of pWDV32 in KI experiments at these two targets (Figure 3), indicating that the performance of pRBKI32N was less robust than that of pRBKI32C. Together, the KI frequencies at 3 targets demonstrated that the DNA tethering strategy with Cas9-Rep robustly increased the HDR-mediated KI frequency.

KI events were inherited by progeny without off-target editing

Next, the inheritance of KI events was analysed for 3 generations. To this end, we selected two monoallelic KI lines of *OsMPK5::HiBiT* (#13 and #16; Figures 4a and S10), which carry a 33 bp-HiBiT-insertion allele (*OsMPK5^{HiBiT}*) and a 2 bp deletion (*OsMPK5^{−2bp}*) allele in the T_0 generation. Among the self-pollinated T_1 plants, the ratios of homozygous and heterozygous KI plants obey the Mendelian ratio (*OsMPK5^{HiBiT/HiBiT}*: *OsMPK5^{−2bp/HiBiT}*:*OsMPK5^{−2bp/−2bp}* = 12:25:11 for line #13 and 3:9:6 for line #16; *chi*-square test, *P* = 0.94 and 0.61 with effect size Cohen's *w* = 0.05 and 0.24 for the T_1 populations of lines #13 and #16, respectively). KI events were also confirmed by HiBiT-based detection of the OsMPK5::HiBiT protein in homozygous and heterozygous KI plants (Figure 4b). We further examined the OsMPK5-HiBiT protein in T_3 homozygous lines, which revealed that the OsMPK5-HiBiT protein was induced by the fungal pathogen *Magnaporthe oryzae* and slightly induced by high-salinity treatment (Figure 4c), as previously reported (Xiong
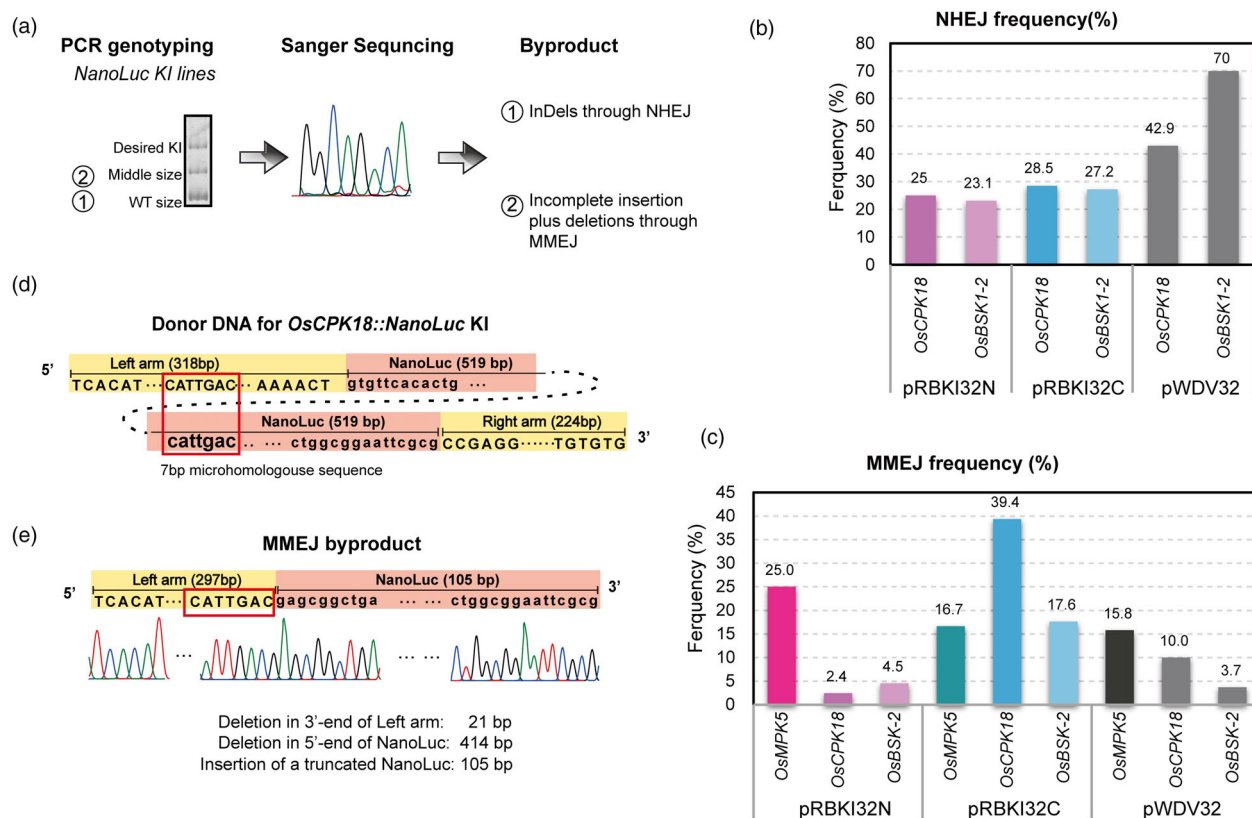
Figure 5 The byproducts in HDR-mediated KI lines. (a) Flowchart showing the detection of byproducts in KI lines. When KI plants were genotyped via PCR, middle-sized products were occasionally detected in the T_0 lines in addition to the desired KI products and wild-type-sized bands. All the PCR products were analysed via Sanger sequencing, and two types of byproducts were found: (1) the wild-type-sized products contained NHEJ-introduced InDels, which are the predominant byproducts, and (2) the middle-sized products contained chromosomal deletions and incomplete insertions from MMEJ, which are the secondary byproducts. (b) Frequencies of NHEJ byproducts detected in the *OsCPK18::NanoLuc* and *OsBSK1-2::NanoLuc* KI lines. (c–e) Analysis of MMEJ byproducts (middle-sized PCR products) in the *OsCPK18::NanoLuc* KI plants. (c) Frequencies of middle-sized byproducts detected in the *OsCPK18::NanoLuc* and *OsBSK1-2::NanoLuc* KI lines. (d, e) An example of MMEJ-mediated incomplete insertion at the OsCPK18 target. The 7-bp microhomologous sequence (red box) is presented in the left arm and *NanoLuc* sequences. (e) Sanger sequencing result showing the undesired KI product in which the 435-bp fragment between the microhomologous sites was deleted.

and Yang, 2003). These data show that the tagged genes in the KI plants were stably transmitted to the progeny.

To assess the frequency of off-target editing or undesired insertions in KI plants, whole-genome sequencing (WGS) was used to analyse 4 homozygous KI plants and 2 homozygous −2 bp mutation plants (Figure 4d, Table S1). The WGS analysis primarily aimed to identify potential random insertions of WDVs and insertions of HiBiT at Cas9 off-target sites. The WGS data revealed the desired KI sequences and a 2 bp deletion at the target site in these 6 plants (Figure S11a). Among the 107 sites that paired with gRNA guides with 1–5 mismatches (Liu et al., 2017), no InDels or KIs were detected in any of the 6 plants (Figure 4e). These data agree with previous reports that Cas9 with properly designed gRNAs has no off-target editing in plants (Tang et al., 2018; Wu et al., 2022). In addition to off-target editing, extrachromosomal WDV DNAs have the potential to be randomly inserted into chromosomes, resulting in unintended integration of WDV sequences into the rice genome. However, WGS did not detect any insertion of the WDV sequence except for the desired HiBiT insertion at the *OsMPK5* target (Figures 4e and S11). The WGS data

demonstrated that the RBKI method generated scar-free KI plants without off-target editing.

The RBKI method reduced NHEJ mutations, and HDR was accompanied by MMEJ byproducts

We next analysed the editing byproducts at the target site (Figure 5a). We first assessed mutagenesis of target genes since NHEJ-introduced InDels are the major byproducts of HDR-mediated KI, prime editing and base editing (Yin and Hu, 2022). To this end, we sequenced the targeting sites of 20 individual T_0 plants of *OsCPK18::NanoLuc* and *OsBSK1-2::NanoLuc*. Interestingly, the NHEJ frequency of the RBKI vectors ranged from 23.8% to 27.5%, which was lower than that of pWDV32 (42%–70%) (Figure 5b), suggesting that NHEJ-mediated edits were reduced with the RBKI method.

During the genotyping of NanoLuc insertion events, we occasionally detected middle-sized PCR products, which were smaller than the desired KI sequences but larger than the WT sequences (Figure 5a), regardless of which vectors were used (Figure 5c). These middle-sized PCR products were detected in 19 *OsCPK18::NanoLuc* plants, 8 *OsMPK5::NanoLuc* plants, and 8

OsBSK1-2::NanoLuc plants (Figures S7–S9). The sequencing results indicate that these middle-sized PCR products in the *OsCPK18::NanoLuc* (Figure 5d,e) and *OsBSK1-2::NanoLuc* (Figure S12) plants represent incomplete insertion coupled with target deletions. In comparison with the desired KI sequence, the middle-sized PCR product of *OsCPK18::NanoLuc* carried a 435 bp deletion from the left homologous arm to *NanoLuc* (Figure 5d,e). Interestingly, the 435-bp fragment was flanked by 7-bp-long microhomologous sequences, suggesting that the deletion event likely resulted from MMEJ (Figure 5d,e). Further characterization of the middle-sized byproducts in *OsBSK1-2::NanoLuc* plants also shows that these incomplete KIs were generated from MMEJ events between the *NanoLuc* and homology arms (Figure S12). We did not observe significant bias in MMEJ among the 3 vectors. However, the occurrence of these MMEJ byproducts was positively correlated with the KI frequencies (Pearson's correlation coefficient $r = 0.67$, $P = 0.0498$), implying that the MMEJ may inherently accompany HDR in DSB repairing. Together, the microhomologous sequences between insertion and targeting genes potentially trigger MMEJ to introduce incomplete insertion byproducts during HDR.

Discussion

Here, we developed the RBKI method for *in vivo* tethering of donor DNA and Cas9, which substantially increased the frequencies of HDR-mediated KI. This method leverages the rolling cycle replication of WDV and the DNA-linking activity of the Rep protein (Figure 1), which enriches the DNA donor to augment HDR frequencies (Figure 2). We obtained monoallelic KI lines at the primary generations for *OsMPK5::HiBiT* and homozygous biallelic KI plants at the T_1 generation (Figure 4). We also show that MMEJ potentially interferes with HDR-mediated precise KI when microhomologous sequences are present between the target and KI sequences. Our study presents a useful toolkit and data to facilitate the application of HDR-mediated genome editing in plants.

This RBKI method was developed on the basis of our hypothesis that the *in situ* availability of donor DNA at the target site is a prerequisite for highly efficient HDR genome editing. This hypothesis was confirmed by ChIP–qPCR and high KI frequencies by targeting 3 rice genes. A notable exception was observed where the Rep-Cas9 fusion failed to enrich donor DNA at the *OsCPK18* target (Figure 2e). We hypothesize that this deficiency may stem from the reduced expression levels of the Rep-Cas9 fusion protein and the suboptimal conformational dynamics between Rep-Cas9 and the WDV replication complex, both of which may compromise efficient DNA-protein tethering at the target site. In contrast, the Cas9-Rep configuration (pRBKI32C vector system) demonstrated robust donor DNA enrichment capacity and significantly enhanced knock-in frequencies in rice. This differential performance suggests that the Cas9-Rep fusion architecture exhibits superior functionality for implementing the RBKI strategy in plant genome editing. Furthermore, the RBKI strategy potentially has additional advantages for the highly efficient HDR events that were observed in this study. First, the WDV replicons tethered to Rep were likely in the form of linear ssDNA (Figure S1), which are better donors than the linear dsDNA in HDR (Shy *et al*., 2023). Second, the RBKI method results in fewer NHEJ mutations (Figure 5b), whereas repression of NHEJ could increase the HDR frequency (Nambiar *et al*., 2022; Qi *et al*., 2013). Third, the overall concentration of circular WDV

replicons was reduced, which may mitigate the cellular toxicity of highly abundant viral DNA (Szewczyk-Roszczenko *et al*., 2025; Yang *et al*., 2022). Finally, the Rep-bridge method tethers donor DNA to the Cas9/gRNA complex *in vivo*, which is different from previous methods. For example, Cas9-VirD fusion tethers DNA and Cas9 during *Agrobacterium* transformation, whereas Cas9-HUH tethers ssODNs *in vitro* (Ali *et al*., 2020; Nagy *et al*., 2022; Tang *et al*., 2023). Since HDR primarily occurs at the S and G2 phases (Yeh *et al*., 2019), persistent *in vivo* tethering of the donor template and Cas9 could increase the likelihood of HDR.

Recent advancements in plant genome engineering have yielded optimized PE tools capable of introducing multinucleotide variations (Cao *et al*., 2024; Zhong *et al*., 2024) and inserting short epitope tags in rice (Li *et al*., 2024; Xu *et al*., 2024). The adaptation of the Duo pegRNA (GRAND) strategy in rice achieved precise integration of short tags with frequencies of up to 70.83% (Xu *et al*., 2024). The MMEJ-based PE strategy also achieved approximately 70% tagging efficiency in rice plants (Li *et al*., 2024). The PE-mediated tag insertion offers operational simplicity and facilitates multiplex editing through co-expression of multiple pegRNAs. However, the insert size is constrained to approximately <100 bp for PEs (Anzalone *et al*., 2019; Xu *et al*., 2024). Moreover, these PE-mediated insertion events are frequently accompanied by additional InDels (Li *et al*., 2024; Xu *et al*., 2024). In contrast, our RBKI strategy demonstrates equivalent efficiency for small tag integration (e.g. HiBiT epitope tagging of *OsMPK5*, Figure 3) while ensuring precise genomic integration without additional InDels at the junction regions. Notably, RBKI enables large-fragment insertion (hundreds of base pairs) with comparable efficiency to short-tag integration. However, larger insertions using RBKI increase susceptibility to cryptic microhomology-mediated repair events, as evidenced by elevated MMEJ byproduct formation (Figures 5c and S12). In all three targets, the frequencies of MMEJ byproducts were lower than the frequencies of desired HDR-mediated KI events (Figures 3 and 5c), implying that MMEJ byproducts would not affect the obtaining of desired KI lines.

Although we achieved highly frequent KI events via the RBKI method, further optimization of this strategy is necessary. Currently, the molecular details of Geminivirus replication, including (1) the functions of Rep, RepA and RepB (Figure S2), (2) the release of DNA from the Rep protein, (3) the restriction of viral replication by the host immune system, and (4) the failure to enrich WDV replicons with Rep-Cas9 at the *OsCPK18* target (Figure 2e), are not fully understood. Unveiling these mechanisms will help to further improve the RBKI method. Additionally, in the RBKI method, the HUH endonuclease domain of the Rep protein responds to the cleavage and rejoining of ssDNA (Figures S1 and S2). A wide repertoire of HUH endonucleases has been identified in all three domains of life and plays key roles in the rolling-circle replication of viruses and plasmids (Ruiz-Maso *et al*., 2015) and in various types of transposition (Chandler *et al*., 2013). These HUH endonucleases could be engineered to persistently conjugate DNAs *in vivo* for diverse organisms. On the other hand, DSB repair pathways, namely, NHEJ, HDR and MMEJ, are poorly understood in plants. For example, we observed variations in KI frequencies among the *OsMPK5*, *OsCPK18*, and *OsBSK1-2* targets, likely due to the site-specific preference of the HDR pathway. The MMEJ byproducts were frequently identified in the *OsCPK18::NanoLuc* KI plants (Figures 5c–e and S8), suggesting that optimizing the design of homologous arms and

insertion sequences is necessary to mitigate MMEJ byproducts in HDR-mediated genome editing. Moreover, the moderate frequency of MMEJ events also suggests that this pathway might be harnessed to shorten the homologous arm sequence. Finally, a recent report demonstrated that an *in vitro* assembled tetra complex, including Cas9, gRNA, ssDNA template, and DNA polymerase, is a functional alternative to prime editing (click editing) (Ferreira da Silva *et al.*, 2024). The *in vivo* DNA-Cas tethering strategy presented in this study can be integrated with this new editing tool for precise manipulation of target gene sequences.

In summary, we present a strategy to tether a DNA donor to Cas9 *in vivo* to increase the efficiency of HDR-mediated genome editing. Given that the Rep protein and other HUH endonucleases are widely available for both plants and animals, we anticipate that this RBKI method will be broadly used in genome editing.

Experimental procedures

Plant materials

Rice (*Oryza sativa* L. *ssp.*) cv. KitaaKe was used in this study for *Agrobacterium tumefaciens*-mediated transformation and protoplast transfection. Rice plants were grown in a greenhouse under 14 h of light at 28 °C and 10 h of dark at 23 °C. The *M. oryzae* infection and salt treatments of the rice plants were performed as described previously (Xie *et al.*, 2014; Xiong and Yang, 2003).

Plasmid vector construction

The engineered WDV replicon (LIR-SIR-Rep-LIR) was synthesized and inserted into pUC57 (GenScript). Notably, the *Bsa*I sites in the WDV sequences were removed via site-directed mutagenesis using a QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent Technologies). The WDV fragment was subsequently amplified via PCR with the oligonucleotide primers 32-HindIII-F and 32-HindIII-R. The resulting PCR product was inserted into the *Hind* III site of the p1300-32 backbone (Xie and Yang, 2013) using the ClonExpress II One Step Cloning Kit (Vazyme), resulting in the p1300-32-WDV plasmid. pWDV32 was constructed by transferring Cas9 from pENTR11 (Xie and Yang, 2013) to p1300-32-WDV via the LR reaction.

To fuse the Rep protein to the C-terminus of Cas9, the stop codon of Cas9 in the pENTR11-Cas9 plasmid was removed via site-directed mutagenesis using a QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent Technologies), resulting in the pENTR11-Cas9-NS plasmid. The *Rep* fragment was amplified via PCR with the oligonucleotides Rep-AflII-F and Rep-SacI-R and then inserted into *Afl*II- and *Sac*I-digested pH-PABE7 (Li *et al.*, 2018), resulting in a *linker-Rep* fusion. The *linker-Rep* fusion fragment was subsequently amplified via PCR with the oligonucleotides Rep-EcoRI-F and Rep-EcoRI-R, and the PCR product was subsequently cloned and inserted into the *Eco*RI site of pENTR11-Cas9-NS via Gibson cloning (New England Biolabs), resulting in pENTR11-Cas9-Rep. Finally, Cas9-Rep in pENTR11 was subcloned and inserted into p1300-32-WDV via the LR reaction, resulting in the pRBKI32C vector.

To construct pRBKI32N, the oligonucleotides Rep-NcoI-F and Rep-32aa-N2-R were used to amplify the *Rep* fragment from pUC57-WDV plasmids, and the oligonucleotides 32aa-Rep-N3-F and 32aa-NcoI-N-R were used to amplify the 32aa-Linker from pH-PABE7 (Li *et al.*, 2018). These two PCR fragments were assembled via overlapped extension PCR. The Rep-Linker fragment was subsequently assembled with *Nco*I-digested

pENTR11-Cas9 via Gibson cloning (New England Biolabs), resulting in the pENTR11-Rep-Cas9 plasmid. Finally, Rep-Cas9 in pENTR11 was transferred to p1300-32-WDV via the LR reaction, resulting in the pRBKI32N vector.

The sequences of the oligonucleotide primers used in vector construction are listed in Table S2. The full sequence of the synthetic WDV replicon is provided in Table S3.

Construction of KI constructs

The procedure used to assemble the plasmids for KI is shown in Figure S13. For the HDR template, the 200–320 bp genomic sequences upstream and downstream of the Cas9 cleavage site were chosen as the left and right homologous arms for HDR, respectively (Figure S5 and Table S3). All backbone vectors generated in this study contained identical cloning sites (Figure S13a) to insert the HDR template and target-specific gRNAs in 5 steps (Figure S13b). First, the left-arm, KI sequence (i.e. NanoLuc), and right-arm fragments were amplified via PCR with specific primers containing appropriate adaptor sequences. Second, the left-arm and KI sequences were assembled via overlapped extension PCR (or Golden Gate assembly). Third, the left-arm-KI fragment from the second step was assembled with the right-arm fragment via overlapped extension PCR (or Golden Gate assembly), resulting in a donor template with the desired 5′- and 3′-adaptor sequences. Fourth, the donor templates were inserted into the *Bam*HI and *Kpn*I sites of pRBKI vectors via Gibson cloning. Fifth, target-specific gRNAs were inserted into the dual BsaI site as we described previously (Xie *et al.*, 2015), resulting in the desired plasmid constructs for KI. The designs of the gRNAs are shown in Figure S5. The primers used to assemble the KI constructs are listed in Table S2. The full sequences of the assembled KI templates for the 3 targets are shown in Table S3.

Protoplast transfection and rice transformation

Rice protoplast preparation and transfection were performed as previously reported (Xie and Yang, 2013). Briefly, 10–20 μg of plasmid DNA was used to transfect 1×10^7 rice protoplasts. A GFP marker was always included to monitor the efficiency of protoplast transfection, and protoplasts whose transfection efficiency was >20% were subjected to subsequent experiments. The protoplasts were harvested after incubation in the dark for 36 h. For stable rice transformation, the plasmid constructs were delivered into the rice cv. KitaaKe via *Agrobacterium*-mediated transformation (TowinBio, Wuhan, China).

Protein extraction and western blotting

Briefly, total proteins were extracted from the protoplasts via an extraction buffer containing 50 mm Tris–HCl (pH 7.4), 150 mm NaCl, 1% (v/v) Triton X-100, 1% (v/v) protease inhibitor cocktail (Sigma–Aldrich, Burlington, MA, USA), and 10% glycerol. For Western blotting detection of Cas9 fusion proteins, 2 μg of total protein from protoplasts was separated via sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and then transferred to a polyvinylidene fluoride (PVDF) membrane. After blocking with 5% (w/v) nonfat milk in Tris-buffered saline with Tween 20 (TBST), the blot was probed with an anti-FLAG antibody (Sigma–Aldrich, 1:1000 dilution) followed by horseradish peroxidase–conjugated anti-mouse IgG (Sigma–Aldrich, 1:10 000 dilution). Finally, the proteins in the blot were detected via a SuperSignal West Femto Maximum Sensitivity Substrate Kit (Life Technologies, Carlsbad, CA, USA) and a Tanon-2500 chemiluminescence imager (Tanon, Shanghai, China).

Analysis of WDV cyclization and copy number

The copy number of WDV replicons in protoplasts was measured via the comparative Ct (threshold cycle) method, which is similar to the qPCR-based assessment of the copy number of transgenes (Bubner and Baldwin, 2004). To calculate the copy number of WDV relative to the rice genome, a single-copy gene (1 copy per rice genome), rice *UBIQUITIN* (*OsUBI*), was used as an endogenous reference. qPCR was performed with a TB Green Premix Ex Taq II (Tli RNaseH Plus) kit (TAKARA Bio) and QuantStudio Prism 3 (Thermo Fisher Scientific). The qPCR primers used are listed in Table S2.

Chromatin immunoprecipitation (ChIP)

For ChIP–qPCR, 5×10^7 rice protoplasts were transfected with KI plasmid constructs derived from the pRBKI32N, pRBKI32C and pWDV32 vectors. After being incubated in WI solution (4 mM MES, pH 5.7, 0.6 M mannitol and 4 mM KCl) for 24 h for ChIP, the protoplasts were collected by centrifugation (300 *g*, 3 min) and resuspended in 1× PBS containing 2% (w/v) formaldehyde. The samples were crosslinked at room temperature for 5 min for ChIP. Then, glycine solution (1 M, pH 2.7) was added to a final concentration of 125 mM to quench the formaldehyde and terminate the crosslinking reaction. The crosslinked protoplasts were washed 2 times with WI solution and resuspended in nuclei isolation buffer containing 20 mM HEPES (pH 8.0), 250 mM sucrose, 1 mM MgCl$_2$, 5 mM KCl, 40% glycerol, 0.25% Triton X-100, 0.1 mM PMSF and 0.1% protease inhibitor cocktail (Sigma–Aldrich). After being incubated in a rocker for 15 min, the nucleic acid was collected by centrifugation at 2300 *g* for 5 min at 4 °C.

For ChIP, the nucleic pellets were mixed with 500 µL of lysis buffer (50 mM Tris–HCl pH 8.0, 10 mM EDTA, 0.5% SDS) and then centrifuged at 16 000 *g* for 10 min at 4 °C. A 10% aliquot of each supernatant was kept as input (10%), and the remaining 450 µL of each supernatant was mixed with 28.7 µL of 5 M NaCl to a final concentration of 300 mM. For immunoprecipitation, 20 µL of anti-FLAG M2 agarose beads (Sigma–Aldrich) were washed 3 times with PBS and then mixed with nucleic extracts. After incubation at 4 °C overnight in a rocker, the beads were washed 2 times with RIPA buffer plus 0.5 M NaCl and then washed 2 times with TE buffer. The protein–DNA complex was eluted with 100 µL of elution buffer (50 mM Tris–HCl pH 8.0, 10 mM EDTA, and 1% SDS) by incubation at 65 °C overnight to reverse crosslinking. Then, 1 µL of RNase A (10 mg/mL) and 2 µL of Protease K (20 mg/mL, Thermo Fisher Scientific) were added to remove RNA and protein by incubation at 37 °C for 30 min. The DNA was extracted with phenol–chloroform–isoamyl alcohol (25:24:1) following ethanol precipitation. Finally, the DNA was dissolved in 20 µL of H$_2$O. For qPCR, WDV-specific primers (SIR-qPCR-F and SIR-qPCR-R) were used to detect WDV replicons in the IP and input samples, with 3 technical replicates. The ChIP-qPCR data were normalized using the Percent Input method (% INPUT) (Solomon *et al.*, 2021). The ChIP signals were calculated with the following equation: % INPUT = $100 \times 2^{((Cq.input-log2(DF))-Cq.IP)}$, Cq.input and Cq.IP represent the quantification cycle of input and IP samples and DF represents the dilution factor. The sequences of the primers used for ChIP–qPCR are listed in Table S2.

Genomic DNA extraction and genotyping

Rice genomic DNA was isolated from protoplasts and leaves via cetyltrimethylammonium bromide (Ding *et al.*, 2018). For all the plants, the quality of the genomic DNA preparations was first examined via PCR amplification of the rice *Actin* gene. The transgene was analysed via Cas9-specific PCR primers. For NanoLuc KI, the genotypes were analysed via a pair of target-specific primers that cross the homologous arms. Notably, the PCR products of NanoLuc KI were separated by native polyacrylamide gel electrophoresis (6%), which provides better resolution than agarose gel electrophoresis does. For the HiBiT tag, the KI events were examined via PCR primers that cover the 5′-junction and 3′-junction regions. Due to the short length of the HiBiT tags, the KI events of *OsMPK5::HiBiT* were further confirmed via nested PCR for T$_0$ plants. In all PCRs, negative controls with WT genomic DNA as the template were included. See Figures S6–S9 for the locations of all the genotyping primers at the target sites and Table S2 for the sequences of the primers used.

HiBiT blotting

HiBiT-tagged proteins were detected via the Nano-Glo HiBiT Blotting System (Promega) following the manufacturer's instructions. Briefly, 20 µg of rice total protein was separated via 10% (w/v) sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and blotted onto a PVDF membrane. The blot was washed with TBST for 4 h at room temperature. Then, the blot was transferred to 1 × Nano-Glo blotting buffer containing LgBiT protein (1:200 dilution) and incubated at room temperature for 2 h. Then, Nano-Glo Luciferase Assay Substrate (1:500 dilution) was added, and the mixture was incubated for 5 min at room temperature. The HiBiT-tagged proteins in the blot were imaged via a Tanon 5200 chemiluminescent imager (Tanon, Shanghai, China).

Whole-genome sequencing analysis

For rice genome sequencing, 1 µg of high-quality genomic DNA was used to prepare the DNA library via the MGIEasy Universal DNA Library Prep Set. The libraries were sequenced via 2 × 150 bp paired-end sequencing via a DNBSEQ-T7 sequencer (MGI Tech, Shenzhen, China). The genome sequences and annotations of *Oryza sativa* L. ssp. *japonica* cv. KitaaKe were downloaded from Phytozome (https://phytozome-next.jgi.doe.gov/). The raw sequencing data were trimmed via fastp (0.23.4) (Chen *et al.*, 2018) to remove adapters, short reads (<50 bp) and low-quality reads with the following parameters: -z 4 -f 5 -t 5 -F 5 -T 5 -5 -W 20 -M 25 -Q -l 50 -c -w 4. The remaining clean reads were mapped to the rice reference genome via BWA-MEM (0.7.17-r1188) (Li and Durbin, 2009) with the following parameters: -M -t 20. The mapped reads were sorted via SAMtools (v1.18). SNPs and indels were identified via BCFtools mpileup (v1.8) with the following parameters: -q 1 -m 2 -F 0.002. The SNPs were filtered as follows via in-house computer scripts: (1) sequence variants identified in more than 2 samples were removed; (2) the remaining variants were further filtered by their average sequencing depth (DP), mapping quality (MQ) and REF/ALT polymorphism probability (QUAL) of each sample. The potential off-target site in the rice genome was predicted via CRISPR-P 2.0 (Liu *et al.*, 2017).

Accession numbers

The sequence data from this study can be found in the NCBI GenBank and MSU rice genome annotations under the following accession numbers: *OsMPK5* (NCBI: LOC4332475; MSU:

LOC_Os03g17700); *OsCPK18* (NCBI: LOC4343002; MSU: LOC_Os07g22710); and *OsBSK1-2* (NCBI: LOC9270011; MSU: LOC_Os10g39670).

Author contributions

K.X. conceived the project and designed the experiments. Z.Z., J.X., C.C. and Y.S. performed the experiments and analysed the data. J.Z. and L.X. analysed the data. J.Z., L.X. and K.X. obtained funding to support the work. K.X. wrote the manuscript with input from all the authors.

Acknowledgements

We thank Ms. Lu Zuo for her help in rice transformation.

Competing interests

The authors have filed a patent application (202310849646.6) on the basis of the data reported in this paper.

Funding

This study was supported by the Science and Technology Innovation 2030-Biological Breeding Major Projects (2023ZD0407403, 2023ZD04062), the National Natural Science Foundation of China (32293243, 31821005), the Fundamental Research Funds for the Central Universities (2662023PY006), the Hainan Yazhou Bay Seed Laboratory and the China National Seed Group (project B23YQ1516).

Data availability statement

The plasmid vectors generated in the current study are available from the corresponding author upon reasonable request. The vector sequences generated or analysed during this study are included in this published article and its supplementary information files. The whole-genome sequencing data that support the findings of this study have been deposited in the Genome Sequence Archive of the National Genomics Data Center, China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (project accession number: PRJCA029228), which are publicly accessible at https://ngdc.cncb.ac.cn/gsa.

References

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.

Ali, Z., Shami, A., Sedeek, K., Kamel, R., Alhabsi, A., Tehseen, M., Hassan, N. *et al*. (2020) Fusion of the Cas9 endonuclease and the VirD2 relaxase facilitates homology-directed repair for precise genome engineering in rice. *Commun. Biol.* **3**, 44.

Anzalone, A.V., Randolph, P.B., Davis, J.R., Sousa, A.A., Koblan, L.W., Levy, J.M., Chen, P.J. *et al*. (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157.

Baltes, N.J., Gil-Humanes, J., Cermak, T., Atkins, P.A. and Voytas, D.F. (2014) DNA replicons for plant genome engineering. *Plant Cell* **26**, 151–163.

Bonnamy, M., Blanc, S. and Michalakis, Y. (2023) Replication mechanisms of circular ssDNA plant viruses and their potential implication in viral gene expression regulation. *MBio* **14**, e0169223.

Bubner, B. and Baldwin, I.T. (2004) Use of real-time PCR for determining copy number and zygosity in transgenic plants. *Plant Cell Rep.* **23**, 263–271.

Cao, Z., Sun, W., Qiao, D., Wang, J., Li, S., Liu, X., Xin, C. *et al*. (2024) PE6c greatly enhances prime editing in transgenic rice plants. *J. Integr. Plant Biol.* **66**, 1864–1870.

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.

Cermak, T., Baltes, N.J., Cegan, R., Zhang, Y. and Voytas, D.F. (2015) High-frequency, precise modification of the tomato genome. *Genome Biol.* **16**, 232.

Chandler, M., de la Cruz, F., Dyda, F., Hickman, A.B., Moncalian, G. and Ton-Hoang, B. (2013) Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. *Nat. Rev. Microbiol.* **11**, 525–538.

Chen, S., Zhou, Y., Chen, Y. and Gu, J. (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics (Oxford, England)* **34**, i884–i890.

Chen, J., Li, S., He, Y., Li, J. and Xia, L. (2022) An update on precision genome editing by homology-directed repair in plants. *Plant Physiol.* **188**, 1780–1794.

Ding, D., Chen, K., Chen, Y., Li, H. and Xie, K. (2018) Engineering introns to express RNA guides for Cas9- and Cpf1-mediated multiplex genome editing. *Mol. Plant* **11**, 542–552.

Dong, O.X. and Ronald, P.C. (2021) Targeted DNA insertion in plants. *Proc. Natl. Acad. Sci. USA* **118**, e2004834117.

Dong, O.X., Yu, S., Jain, R., Zhang, N., Duong, P.Q., Butler, C., Li, Y. *et al*. (2020) Marker-free carotenoid-enriched rice generated through targeted gene insertion using CRISPR-Cas9. *Nat. Commun.* **11**, 1178.

Endo, M., Mikami, M. and Toki, S. (2016) Biallelic gene targeting in rice. *Plant Physiol.* **170**, 667–677.

Ferreira da Silva, J., Tou, C.J., King, E.M., Eller, M.L., Rufino-Ramos, D., Ma, L., Cromwell, C.R. *et al*. (2024) Click editing enables programmable genome writing using DNA polymerases and HUH endonucleases. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-41024-02324-x

Gao, C. (2021) Genome engineering for crop improvement and future agriculture. *Cell* **184**, 1621–1635.

Gil-Humanes, J., Wang, Y., Liang, Z., Shan, Q., Ozuna, C.V., Sanchez-Leon, S., Baltes, N.J. *et al*. (2017) High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. *Plant J.* **89**, 1251–1262.

Hall, M.P., Unch, J., Binkowski, B.F., Valley, M.P., Butler, B.L., Wood, M.G., Otto, P. *et al*. (2012) Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chem. Biol.* **7**, 1848–1857.

Hanley-Bowdoin, L., Bejarano, E.R., Robertson, D. and Mansoor, S. (2013) Geminiviruses: masters at redirecting and reprogramming plant processes. *Nat. Rev. Microbiol.* **11**, 777–788.

Laufs, J., Traut, W., Heyraud, F., Matzeit, V., Rogers, S.G., Schell, J. and Gronenborn, B. (1995) In vitro cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. *Proc. Natl. Acad. Sci. USA* **92**, 3879–3883.

Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760.

Li, C., Zong, Y., Wang, Y., Jin, S., Zhang, D., Song, Q., Zhang, R. *et al*. (2018) Expanded base editing in rice and wheat using a Cas9-adenosine deaminase fusion. *Genome Biol.* **19**, 59.

Li, H., Wu, C., Du, M., Chen, Y., Hou, X., Yang, Y. and Xie, K. (2021) A versatile nanoluciferase toolkit and optimized in-gel detection method for protein analysis in plants. *Mol. Breed.* **41**, 13.

Li, X., Zhang, S., Wang, C., Ren, B., Yan, F., Li, S., Spetz, C. *et al*. (2024) Efficient in situ epitope tagging of rice genes by nuclease-mediated prime editing. *Plant Cell* **37**(2), koae316.

Ling, X., Xie, B., Gao, X., Chang, L., Zheng, W., Chen, H., Huang, Y. *et al*. (2020) Improving the efficiency of precise genome editing with site-specific Cas9-oligonucleotide conjugates. *Sci. Adv.* **6**, eaaz0051.

Liu, H., Ding, Y., Zhou, Y., Jin, W., Xie, K. and Chen, L.L. (2017) CRISPR-P 2.0: an improved CRISPR-Cas9 tool for genome editing in plants. *Mol. Plant* **10**, 530–532.

Liu, P., Panda, K., Edwards, S.A., Swanson, R., Yi, H., Pandesha, P., Hung, Y.H. *et al*. (2024) Transposase-assisted target-site integration for efficient plant genome engineering. *Nature* **631**, 593–600.

Lovendahl, K.N., Hayward, A.N. and Gordon, W.R. (2017) Sequence-directed covalent protein-DNA linkages in a single step using HUH-tags. *J. Am. Chem. Soc.* **139**, 7030–7035.

Lu, Y., Tian, Y., Shen, R., Yao, Q., Wang, M., Chen, M., Dong, J. *et al*. (2020) Targeted, efficient sequence insertion and replacement in rice. *Nat. Biotechnol.* **38**, 1402–1407.

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.

Matzeit, V., Schaefer, S., Kammann, M., Schalk, H.J., Schell, J. and Gronenborn, B. (1991) Wheat dwarf virus vectors replicate and express foreign genes in cells of monocotyledonous plants. *Plant Cell* **3**, 247–258.

Miki, D., Zhang, W., Zeng, W., Feng, Z. and Zhu, J.K. (2018) CRISPR/Cas9-mediated gene targeting in *Arabidopsis* using sequential transformation. *Nat. Commun.* **9**, 1967.

Missich, R., Ramirez-Parra, E. and Gutierrez, C. (2000) Relationship of oligomerization to DNA binding of Wheat dwarf virus RepA and Rep proteins. *Virology* **273**, 178–188.

Nagy, E.D., Kuehn, R., Wang, D., Shrawat, A., Duda, D.M., Groat, J.R., Yang, P. *et al*. (2022) Site-directed integration of exogenous DNA into the soybean genome by LbCas12a fused to a plant viral HUH endonuclease. *Plant J.* **111**, 905–916.

Nambiar, T.S., Baudrier, L., Billon, P. and Ciccia, A. (2022) CRISPR-based genome editing through the lens of DNA repair. *Mol. Cell* **82**, 348–388.

Qi, Y., Zhang, Y., Zhang, F., Baller, J.A., Cleland, S.C., Ryu, Y., Starker, C.G. *et al*. (2013) Increasing frequencies of site-specific mutagenesis and gene targeting in Arabidopsis by manipulating DNA repair pathways. *Genome Res.* **23**, 547–554.

Renkawitz, J., Lademann, C.A. and Jentsch, S. (2014) Mechanisms and principles of homology search during recombination. *Nat. Rev. Mol. Cell Biol.* **15**, 369–383.

Ruiz-Maso, J.A., Macho, N.C., Bordanaba-Ruiseco, L., Espinosa, M., Coll, M. and Del Solar, G. (2015) Plasmid rolling-circle replication. *Microbiol. Spectr.* **3**, PLAS-0035-2014.

Savic, N., Ringnalda, F.C., Lindsay, H., Berk, C., Bargsten, K., Li, Y., Neri, D. *et al*. (2018) Covalent linkage of the DNA repair template to the CRISPR-Cas9 nuclease enhances homology-directed repair. *elife* **7**, e33761.

Schalk, H.J., Matzeit, V., Schiller, B., Schell, J. and Gronenborn, B. (1989) Wheat dwarf virus, a geminivirus of graminaceous plants needs splicing for replication. *EMBO J.* **8**, 359–364.

Schwinn, M.K., Machleidt, T., Zimmerman, K., Eggers, C.T., Dixon, A.S., Hurst, R., Hall, M.P. *et al*. (2018) CRISPR-mediated tagging of endogenous proteins with a luminescent peptide. *ACS Chem. Biol.* **13**, 467–474.

Shy, B.R., Vykunta, V.S., Ha, A., Talbot, A., Roth, T.L., Nguyen, D.N., Pfeifer, W.G. *et al*. (2023) High-yield genome engineering in primary cells using a hybrid ssDNA repair template and small-molecule cocktails. *Nat. Biotechnol.* **41**, 521–531.

Solomon, E.R., Caldwell, K.K. and Allan, A.M. (2021) A novel method for the normalization of ChIP-qPCR data. *MethodsX* **8**, 101504.

Sun, Y., Zhang, X., Wu, C., He, Y., Ma, Y., Hou, H., Guo, X. *et al*. (2016) Engineering herbicide-resistant rice plants through CRISPR/Cas9-mediated homologous recombination of *Acetolactate synthase. Mol. Plant* **9**, 628–631.

Sun, C., Lei, Y., Li, B., Gao, Q., Li, Y., Cao, W., Yang, C. *et al*. (2024) Precise integration of large DNA sequences in plant genomes using PrimeRoot editors. *Nat. Biotechnol.* **42**, 316–327.

Szewczyk-Roszczenko, O., Roszczenko, P., Vassetzky, Y. and Sjakste, N. (2025) Genotoxic consequences of viral infections. *NPJ Viruses* **3**, 5.

Tan, J., Zhao, Y., Wang, B., Hao, Y., Wang, Y., Li, Y., Luo, W. *et al*. (2020) Efficient CRISPR/Cas9-based plant genomic fragment deletions by microhomology-mediated end joining. *Plant Biotechnol. J.* **18**, 2161–2163.

Tang, X., Liu, G., Zhou, J., Ren, Q., You, Q., Tian, L., Xin, X. *et al*. (2018) A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. *Genome Biol.* **19**, 84.

Tang, Y., Zhang, Z., Yang, Z. and Wu, J. (2023) CRISPR/Cas9 and *Agrobacterium tumefaciens* virulence proteins synergistically increase efficiency of precise genome editing via homology directed repair in plants. *J. Exp. Bot.* **74**, 3518–3530.

Timmermans, M.C., Das, O.P. and Messing, J. (1992) *Trans* replication and high copy numbers of wheat dwarf virus vectors in maize cells. *Nucleic Acids Res.* **20**, 4047–4054.

Tompkins, K.J., Houtti, M., Litzau, L.A., Aird, E.J., Everett, B.A., Nelson, A.T., Pornschloegl, L. *et al*. (2021) Molecular underpinnings of ssDNA specificity by Rep HUH-endonucleases and implications for HUH-tag multiplexing and engineering. *Nucleic Acids Res.* **49**, 1046–1064.

Wang, J.Y. and Doudna, J.A. (2023) CRISPR technology: a decade of genome editing is only the beginning. *Science* **379**, eadd8643.

Wang, M., Lu, Y., Botella, J.R., Mao, Y., Hua, K. and Zhu, J.K. (2017) Gene targeting by homology-directed repair in rice using a geminivirus-based CRISPR/Cas9 system. *Mol. Plant* **10**, 1007–1010.

Wu, Y., Ren, Q., Zhong, Z., Liu, G., Han, Y., Bao, Y., Liu, L. *et al*. (2022) Genome-wide analyses of PAM-relaxed Cas9 genome editors reveal substantial off-target effects by ABE8e in rice. *Plant Biotechnol. J.* **20**, 1670–1682.

Xie, K. and Yang, Y. (2013) RNA-guided genome editing in plants using a CRISPR-Cas system. *Mol. Plant* **6**, 1975–1983.

Xie, K., Chen, J., Wang, Q. and Yang, Y. (2014) Direct phosphorylation and activation of a mitogen-activated protein kinase by a calcium-dependent protein kinase in rice. *Plant Cell* **26**, 3077–3089.

Xie, K., Minkenberg, B. and Yang, Y. (2015) Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl. Acad. Sci. USA* **112**, 3570–3575.

Xiong, L. and Yang, Y. (2003) Disease resistance and abiotic stress tolerance in rice are inversely modulated by an abscisic acid-inducible mitogen-activated protein kinase. *Plant Cell* **15**, 745–759.

Xu, R., Ma, C., Sheng, J., Zhu, J., Wang, D., Liu, X., Wang, Q. *et al*. (2024) Engineering PE6 prime editors to efficiently insert tags in rice. *Plant Biotechnol. J.* **22**, 3383–3385.

Yang, L., Jia, R., Ge, T., Ge, S., Zhuang, A., Chai, P. and Fan, X. (2022) Extrachromosomal circular DNA: biogenesis, structure, functions and diseases. *Signal Transduct. Target. Ther.* **7**, 342.

Yeh, C.D., Richardson, C.D. and Corn, J.E. (2019) Advances in genome editing through control of DNA repair pathways. *Nat. Cell Biol.* **21**, 1468–1478.

Yin, J. and Hu, J. (2022) The origin of unwanted editing byproducts in gene editing. *Acta Biochim. Biophys. Sin. Shanghai* **54**, 767–781.

Zhong, Z., Fan, T., He, Y., Liu, S., Zheng, X., Xu, Y., Ren, J. *et al*. (2024) An improved plant prime editor for efficient generation of multiple-nucleotide variations and structural variations in rice. *Plant Commun.* **5**, 100976.

Zhou, X., Zhao, Y., Ni, P., Ni, Z., Sun, Q. and Zong, Y. (2023) CRISPR-mediated acceleration of wheat improvement: advances and perspectives. *J. Genet. Genomics* **50**, 815–834.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 Schematics illustrating WDV replication and Rep-DNA conjugation.

Figure S2 Alternative splicing of the *Rep-Cas9* gene.

Figure S3 NHEJ-induced InDels at the *OsMPK5* target using pRBKI vectors.

Figure S4 Comparisons of the copy numbers (CVs) of extrachromosomal WDV replicons in rice protoplasts transfected with 3 vectors.

Figure S5 Schematics showing the design of the donor DNA template and gRNAs for KI at the *OsMPK5*, *OsCPK18*, and *OsBSK1-2* loci.

Figure S6 Genotyping results for the *OsMPK5::HiBiT* T_0 plants.

Figure S7 Genotyping results for the *OsMPK5::NanoLuc* T_0 plants.

Figure S8 Genotyping results for the *OsCPK18::NanoLuc* T_0 plants.

Figure S9 Genotyping results for the *OsBSK1-2::NanoLuc* T_0 plants.

Figure S10 Segregation of KI events for 2 *OsMPK5::HiBiT* lines.

Figure S11 Whole-genome sequencing analysis of 6 T_1 plants derived from *OsMPK5::HiBiT* lines #13 and #16.

Figure S12 MMEJ-mediated incomplete KI at the *OsBSK1-2* target.

Figure S13 Cloning procedure for pRBKI vector construction.

Table S1 Summary of whole-genome sequencing data.

Table S2 Sequences of primers used in this study.

Table S3 Sequences of WDV and donor DNA template for KI.