Plant Biotechnology Journal
Open Access

Plant Biotechnology Journal (2024) **22**, pp. 2488–2503 doi: 10.1111/pbi.14363

Expanding plant genome editing scope and profiles with CRISPR-FrCas9 systems targeting palindromic TA sites

Yao He[1,2,†], Yangshuo Han[2,†], Yanqin Ma[1], Shishi Liu[1], Tingting Fan[1], Yanling Liang[1], Xu Tang[2], Xuelian Zheng[1,2], Yuechao Wu[3,4], Tao Zhang[3,4,*], Yiping Qi[5,6,*] and Yong Zhang[1,2,*]

[1]*Department of Biotechnology, School of Life Sciences and Technology, Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu, China*

[2]*Chongqing Key Laboratory of Plant Resource Conservation and Germplasm Innovation, Integrative Science Center of Germplasm Creation in Western China (Chongqing) Science City, School of Life Sciences, Southwest University, Chongqing, China*

[3]*Jiangsu Key Laboratory of Crop Genomics and Molecular Breeding/Zhongshan Biological Breeding Laboratory/Key Laboratory of Plant Functional Genomics of the Ministry of Education, Agricultural College of Yangzhou University, Yangzhou University, Yangzhou, China*

[4]*Jiangsu Co-Innovation Center for Modern Production Technology of Grain Crops/Jiangsu Key Laboratory of Crop Genetics and Physiology, Yangzhou University, Yangzhou, China*

[5]*Department of Plant Science and Landscape Architecture, University of Maryland, College Park, Maryland, USA*

[6]*Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, Maryland, USA*

Received 16 December 2023;
revised 14 March 2024;
accepted 2 April 2024.
**Correspondence* (Tel 86-18780156670; email zhangyong916@uestc.edu.cn (YZ), Tel 01-301-405-7682; email yiping@umd.edu (YQ), Tel 86-514-87977229; email zhangtao@yzu.edu.cn (TZ))
[†]These authors have contributed equally to this work.

Keywords: CRISPR-FrCas9, TA PAM requirement, TREX2, base editor, microRNA knockout, off-target analysis.

Summary

CRISPR-Cas9 is widely used for genome editing, but its PAM sequence requirements limit its efficiency. In this study, we explore *Faecalibaculum rodentium* Cas9 (FrCas9) for plant genome editing, especially in rice. FrCas9 recognizes a concise 5′-NNTA-3′ PAM, targeting more abundant palindromic TA sites in plant genomes than the 5′-NGG-3′ PAM sites of the most popular SpCas9. FrCas9 shows cleavage activities at all tested 5′-NNTA-3′ PAM sites with editing outcomes sharing the same characteristics of a typical CRISPR-Cas9 system. FrCas9 induces high-efficiency targeted mutagenesis in stable rice lines, readily generating biallelic mutants with expected phenotypes. We augment FrCas9's ability to generate larger deletions through fusion with the exonuclease, TREX2. TREX2-FrCas9 generates much larger deletions than FrCas9 without compromise in editing efficiency. We demonstrate TREX2-FrCas9 as an efficient tool for genetic knockout of a microRNA gene. Furthermore, FrCas9-derived cytosine base editors (CBEs) and adenine base editors (ABE) are developed to produce targeted C-to-T and A-to-G base edits in rice plants. Whole-genome sequencing-based off-target analysis suggests that FrCas9 is a highly specific nuclease. Expression of TREX2-FrCas9 in plants, however, causes detectable guide RNA-independent off-target mutations, mostly as single nucleotide variants (SNVs). Together, we have established an efficient CRISPR-FrCas9 system for targeted mutagenesis, large deletions, C-to-T base editing, and A-to-G base editing in plants. The simple palindromic TA motif in the PAM makes the CRISPR-FrCas9 system a promising tool for genome editing in plants with an expanded targeting scope.

Introduction

Clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 9 (Cas9) from the type II CRISPR-Cas bacterial adaptive immune system has been engineered as an efficient genome editing platform for a wide range of organisms (Sander and Joung, 2014), including plants (Tang and Zhang, 2023; Zhong *et al.*, 2019). The target site recognition of Cas9 is programmed by a chimeric single guide RNA (sgRNA) that encodes a sequence complementary to the target protospacer and by Cas9 that recognizes a short protospacer adjacent motif (PAM) (Jinek *et al.*, 2012). PAM specificity is a critical factor in designing efficient and precise genome editing tools (Jiang *et al.*, 2013). While the widely used *Streptococcus pyogenes* Cas9 (SpCas9) primarily recognizes the 5′-NGG-3′ PAM sequence (Jiang *et al.*, 2013), researchers have explored Cas9 orthologues from various bacterial species to broaden the range of targetable genomic sites. For example, *Staphylococcus aureus* Cas9 (SaCas9), with a 5′-NNGRRT-3′ PAM, has a smaller size than SpCas9 (Ran *et al.*, 2015), making it advantageous for delivery into cells using viral vectors with limited cargo capacity. *Streptococcus thermophilus* Cas9 (St1Cas9), with a 5′-NNAGAAW-3′ PAM, has been investigated for its potential in genome engineering (Kleinstiver *et al.*, 2015). *Neisseria meningitidis* Cas9 (NmCas9) has a relatively longer PAM of 5′-NNNNGMTT-3′, expanding the targeting options in genome engineering (Zhu *et al.*, 2019). However, these Cas9 orthologs use more complex and longer PAMs than SpCas9, which makes them less popular tools in genome engineering.

Efforts to expand the targeting scope of CRISPR-Cas9 systems have led to the development of engineered Cas9 variants with altered or relaxed PAM requirements for genome engineering in eukaryotes, including plants. Several notable examples include SpCas9-VQR (Hu *et al.*, 2016), SpCas9-EQR (Hu *et al.*, 2016), SpCas9-VRER (Qin *et al.*, 2019), iSpyMacCas9 (Chatterjee *et al.*, 2020; Sretenovic *et al.*, 2021), Cas9-NG (Nishimasu *et al.*, 2018;

Zhong *et al*., 2019), and PAM-less SpRY (Ren *et al*., 2021c; Walton *et al*., 2020). However, the alteration in PAM specificity often comes at the cost of reduced genome editing efficiency and increased off-target effects. For instance, SpRY has been observed to exhibit lower genome editing activity compared to the wild-type SpCas9, partly due to its self-editing nature when delivered as DNA (Sretenovic *et al*., 2023b; Walton *et al*., 2020). As an enhanced and hybrid variant of SpCas9 and SmacCas9, iSpyMacCas9 was designed to recognize a more permissive 5′-NNA-3′ PAM sequence, expanding the range of targetable genomic sites (Chatterjee *et al*., 2020). Nevertheless, our study in plants showed iSpyMac9 prefers 5′-NNAR-3′ PAM sites and is an overall less robust nuclease compared to SpCas9 (Sretenovic *et al*., 2021). Therefore, researchers need to assess the trade-off between PAM flexibility and editing efficiency when selecting Cas9 variants for genome editing.

Recently, researchers identified a Type II-A Cas9 ortholog, FrCas9, derived from *Faecalibaculum rodentium* (Cui *et al*., 2022), which exhibited unique biochemical characteristics not previously reported in CRISPR systems. FrCas9 conferred efficient genome editing at 5′-NNTA-3′ PAM sites in human cells. The simple and palindromic nature of the core PAM motif 'TA' increases target site densities in various organisms. Importantly, FrCas9 showed comparable editing efficiency and specificity to the well-known SpCas9 in human cells (Cui *et al*., 2022). These findings highlight the potential of FrCas9 as a versatile and robust genome editing tool in other organisms. In this study, we aimed to develop an efficient FrCas9-based genome editing system in plants. As with many such exploratory studies, we used rice as our test organism. We started with evaluating the sgRNA scaffold's high performance, followed by a comprehensive assessment of FrCas9's PAM requirements and editing robustness across many target sites. To make FrCas9 an efficient tool for inducing larger deletions, a characteristic of CRISPR-Cas12 systems (Ming *et al*., 2020; Tang *et al*., 2017), we generated a TREX2-FrCas9 fusion, which was successfully used to efficiently generate miRNA gene knockout in rice. We also developed FrCas9-based cytosine base editors (CBEs) and adenine base editors (ABEs). Finally, we showed with whole genome sequencing (WGS) that FrCas9, albeit its simple PAM, is a highly specific nuclease for genome editing in plants.

Results

Characterization of CRISPR-FrCas9 for genome editing in plants

Phylogenetic analysis showed that FrCas9 is closely related to SpCas9 (Jiang *et al*., 2013), ScCas9 (Xu *et al*., 2020), and LrCas9 (Zhong *et al*., 2023), all belonging to the Type II-A Cas9 group (Zhong *et al*., 2023). FrCas9 has shown efficient editing capability for target DNA sequences with a 5′-NNTA-3′ PAM in both prokaryotic and human genomes (Cui *et al*., 2022), suggesting this PAM may also be preferred in plants. FrCas9 has a similar size to SpCas9, and both Cas9 systems share high similarity in protein domains and gene structures (Figure S1). However, FrCas9 differs significantly from SpCas9 in that it has seven active residues in RuvC and HNH domains (Cui *et al*., 2022), whereas SpCas9 only has two active residues in RuvC and HNH domains (Figure 1a). FrCas9 has the catalytic centres of RuvCI (D20) and HNH (H877), corresponding to D10 and H840 residues in SpCas9 (Figure 1a). SpCas9 nickases based on D10A and H840A mutations have been used, respectively, to develop efficient base editors (Gaudelli *et al*., 2017; Komor *et al*., 2016) and prime editors (Anzalone *et al*., 2019). Based on SWISS-MODEL prediction of protein structures, the HNH domain is closely associated with the REC domain in FrCas9, while the HNH domain is partly associated with the REC domain and RuvCIII domains in SpCas9 (Figure 1b). Compared to SpCas9, the palindromic 5′-NNTA-3′ PAM of FrCas9 would increase the densities of sgRNA distributions (Figure 1c). Indeed, an *in silico* analysis showed that more target sites can be designed with FrCas9 than SpCas9 in the coding and non-coding regions of the rice genome (Figure 1d). Hence, FrCas9 greatly complements SpCas9 for genome engineering in plants.

In a recent study (Cui *et al*., 2022), it was shown that the FrCas9 sgRNA consisting of truncated crRNA and tracrRNA achieved the best editing efficiency in human cells. To obtain the best sgRNA scaffold for efficient FrCas9 genome editing in plants, we tested three sgRNA scaffolds of FrCas9, including a long tetraloop (sgRNA V01) and two truncated versions (sgRNA V02 and sgRNA V03) with a shorter 3′ terminus of crRNA and a 5′ terminus of tracrRNA (Figure 1e). We tested the three versions of sgRNAs by using 22 bp spacers at four independent 5′-NNTA-3′ PAM sites. Based on next-generation sequencing (NGS) of PCR amplicons, our rice protoplast data showed that the sgRNA V02 worked well in most sites (Figure 1f), consistent with the data in human cells (Cui *et al*., 2022).

To assess whether FrCas9 can robustly edit 5′-NNTA-3′ PAM sites in the rice genome, we targeted 16 endogenous sites in rice protoplasts with 22 bp spacers and the sgRNA V02 scaffold. Genome editing efficiency at these sites was measured by NGS of PCR amplicons. Remarkably, the data showed detectable editing activity by FrCas9 at all 16 target sites, with editing efficiencies ranging from 2.7% to 53.9% (Figure 1g). The NGS data allowed us to investigate editing profiles by FrCas9 at all sites. Firstly, the insertion and deletion (indel) proportions were calculated, and interestingly, FrCas9 caused more insertions than deletions at most target sites (Figure 1h). About deletions, results showed high-frequency deletions around the Cas9 cleavage site (Figure 1i), and a 1 bp deletion is the predominant deletion size (Figure 1j). To our slight surprise, the results revealed that the main frequency of insertions occurred around the PAM (Figure 1k), with 1 bp deletions being the predominant size of insertion (Figure 1l). The above data showed that CRISPR-FrCas9 is an efficient genome editing system for introducing small indels in plant cells.

High-frequency genome editing by CRISRP-FrCas9 in stable rice plants

We next tested whether CRISPR-FrCas9 could efficiently generate edits in stable rice lines. Six CRISPR-FrCas9 constructs were made to target *OsEPFL9, OsGn1a, OsGS3, OsGBSSI, OsROC5*, and *OsDEP1*, respectively (Figure 2a). Encouragingly, mutated T0 generation plants can be obtained at all six sites, with mutation efficiency ranging from 10% to 85.7%. Out of the six sites we tested, five had biallelic mutations, with biallelic editing efficiency ranging from 35% to 80.7% (Figure 2b). Analysis of mutant plants showed that the mutations were mainly 1 bp insertions, with the remaining mutations being deletions of one or multiple base pairs (Figures S2–S4), which was consistent with the editing profile in rice protoplasts (Figure 1h). At the *OsGn1a*-sgR01 target site, FrCas9 generated 84.6% (22 out of 26 T0 lines) indel frequency (Figure 2b). Among the T0 lines, 80.7% (21 out of 26)

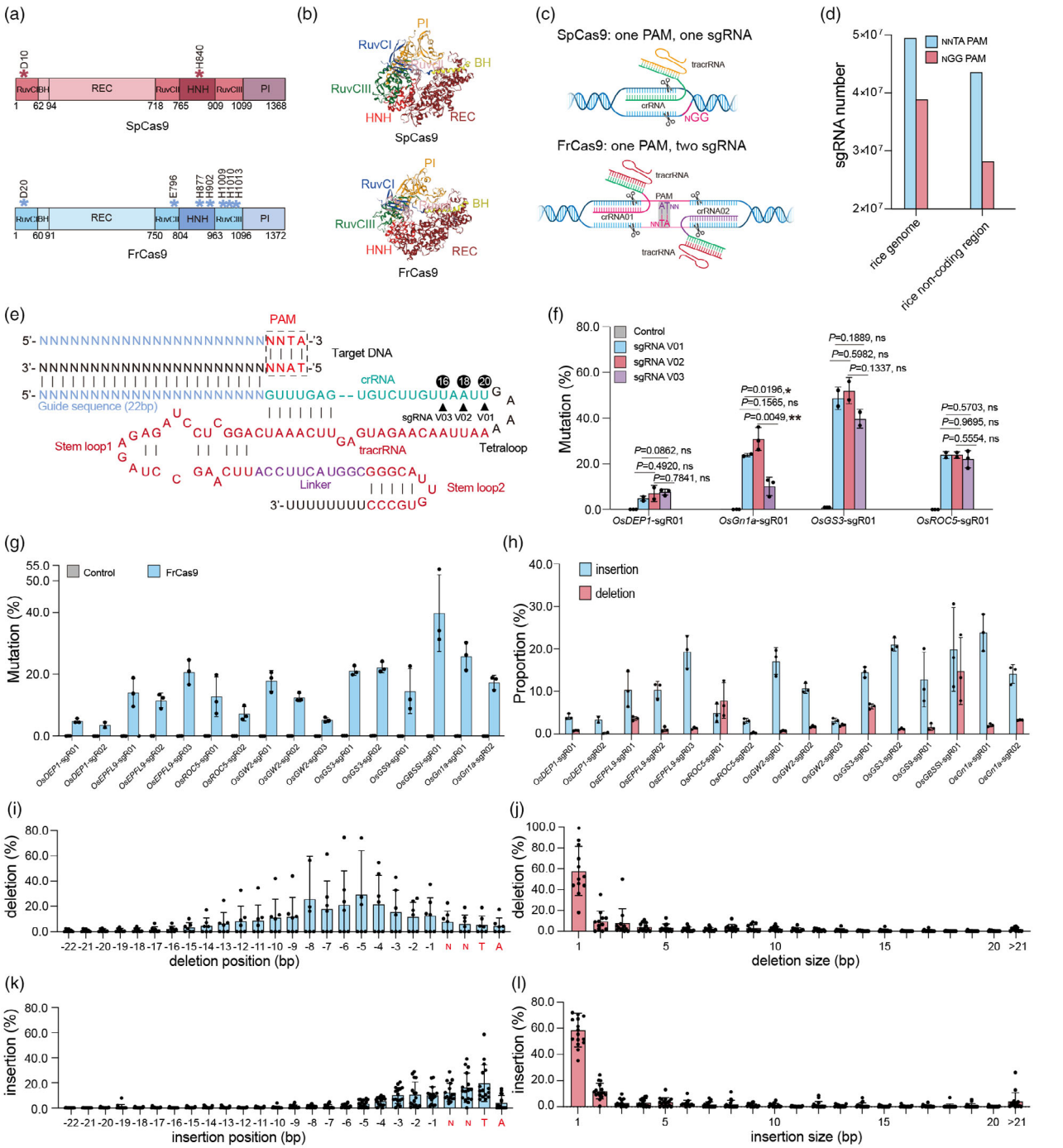
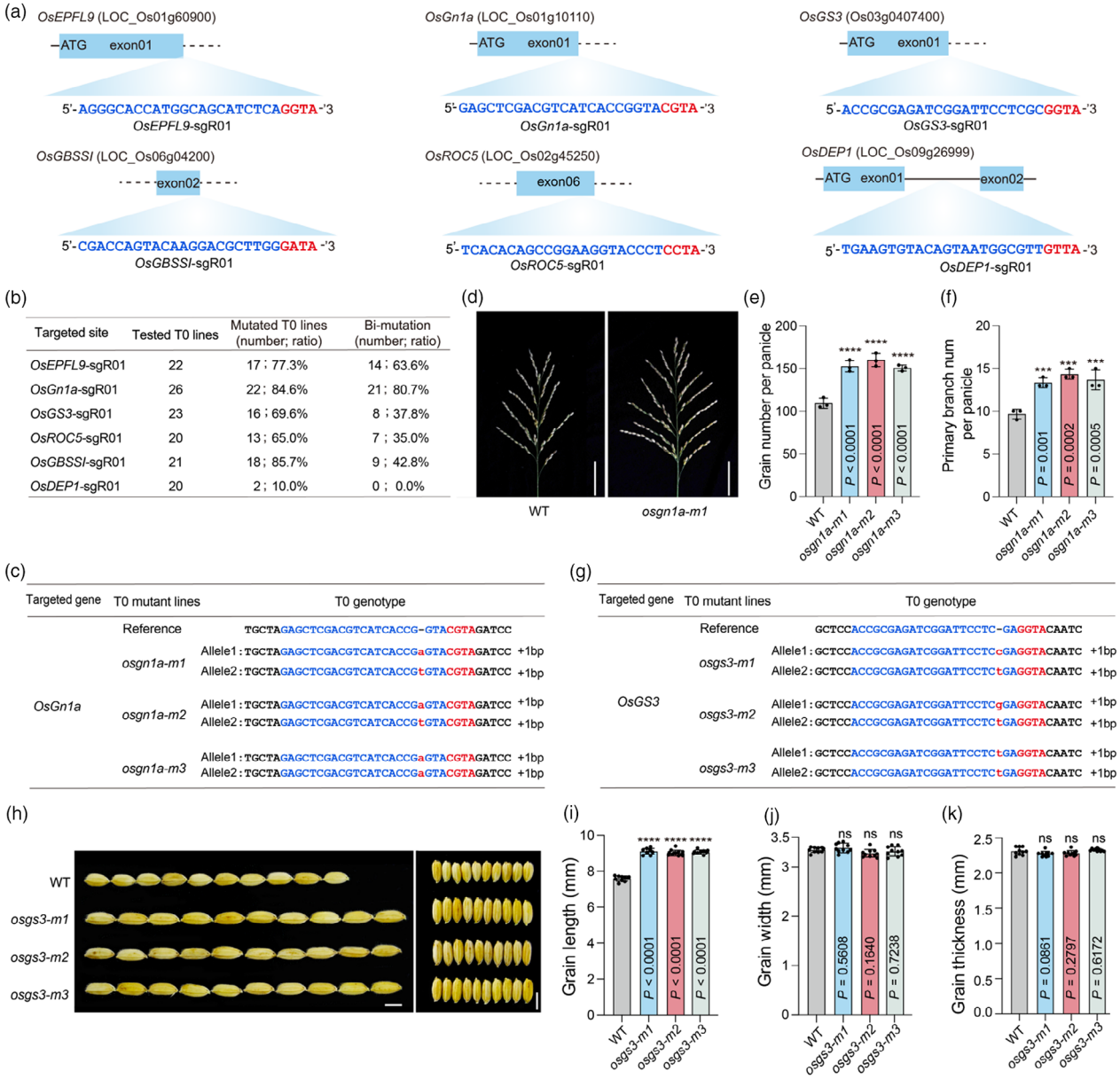


Figure 1 Characterization of CRISPR-FrCas9 for genome editing in plants. (a–d) The comparison of FrCas9 and SpCas9. (a) The domains of FrCas9 and SpCas9 nucleases and the catalytic residues of RuvC and HNH domains of these two nucleases are represented by pentagram. BH Bridge helix domain, PI PAM-interacting domain. (b) Predicted protein structures of FrCas9 and SpCas9 by SWISS-MODEL. (c, d) Characteristic comparisons of the two CRISPR systems support CRISPR–FrCas9 as a non-coding region editing tool. TracrRNA, trans-activating CRISPR RNA. (e) The schematic representation of the sgRNA:target DNA complex. The protospacer (spacer) was labelled in light blue, the crRNA was labelled in green, the RNA linker GAAA was labelled in black, and the tracrRNA was labelled in red and purple. The different tetraloop lengths were pointed out by triangles and named sgRNA V01, sgRNA V02, and sgRNA V03. The lengths of complementary regions were noted above the crRNA sequence with circles. (f) The efficiency of sgRNAs with three different scaffolds of FrCas9 was assayed by target amplicon sequencing. Data are presented as mean values ± SD (n = 2 or 3 biologically independent replicates). (g) The mutation rates of 16 endogenous sites in rice protoplasts by FrCas9. Each dot represents a biological replicate of an independent experiment. Each target contains three biological replicates (n = 3). Data are presented as mean values ± SD. (h) The insertion and deletion proportions by FrCas9 at 16 rice loci. Each dot represents a biological replicate. Each assay contains three independent experiments (n = 3). Data are presented as mean values ± SD. (i) The deletion positions of all sites in (g). Each dot represents a biological replicate. Data are presented as mean values ± SD. (j) The deletion sizes of all sites (g). Each dot represents a biological replicate. Data are presented as mean values ± SD. (k) The insertion positions of all sites (g). Each dot represents a biological replicate. Data are presented as mean values ± SD. (l) The insertion sizes of all sites in (g). Each dot represents a biological replicate. Data are presented as mean values ± SD.



Targeted site	Tested T0 lines	Mutated T0 lines (number; ratio)	Bi-mutation (number; ratio)
OsEPFL9-sgR01	22	17 ; 77.3%	14 ; 63.6%
OsGn1a-sgR01	26	22 ; 84.6%	21 ; 80.7%
OsGS3-sgR01	23	16 ; 69.6%	8 ; 37.8%
OsROC5-sgR01	20	13 ; 65.0%	7 ; 35.0%
OsGBSSI-sgR01	21	18 ; 85.7%	9 ; 42.8%
OsDEP1-sgR01	20	2 ; 10.0%	0 ; 0.0%

Figure 2 CRISPR-FrCas9 confers efficient genome editing in stable T0 rice plants. (a) Schematics of the target sites in *OsEPFL9*, *OsGn1a*, *OsGS3*, *OsGBSSI*, *OsROC5*, and *OsDEP1*. (b) A table summarizing the genome editing efficiency of FrCas9 in stable rice lines at six target sites. Bi-mutation, biallelic mutations. (c) Genotypes of three biallelic mutants derived from FrCas9-mediated editing in *OsGn1a* target site. (d) Panicle type of *OsGn1a* mutants and WT. Bar = 5 cm. (e) Grain number per panicle of different *OsGn1a* mutants and WT. (f) Primary branch number per panicle of different *OsGn1a* mutants and WT. Error bars represent the standard deviations of three biological replicates. Data are means ± SD. Statistical significance is indicated by asterisks. ns, $P > 0.05$; ***$P < 0.001$; ****$P < 0.0001$, one-way ANOVA, Dunnett's test. (g) Genotypes of three biallelic mutants derived from FrCas9-mediated editing at the *OsGS3* target site. (h) Grain size of *OsGS3* mutants and WT. Bar = 0.5 cm. (i–k) Grain length (i), grain width (j), and grain thickness (k) of different *OsGS3* mutants and WT. Error bars represent the standard deviations of 10 biological replicates. Data are means ± SD. Statistical significance is indicated by asterisks. ns, $P > 0.05$; ***$P < 0.001$; ****$P < 0.0001$, one-way ANOVA, Dunnett's test.

were biallelically edited, such as *osgn1a-m1*, *osgn1a-m2*, and *osgn1a-m3* lines (Figure 2c). *OsGn1a* biallelic mutants by FrCas9 all showed increased grain number per panicle and primary branch number per panicle (Figure 2d–f), which are anticipated phenotypes of *OsGn1a* knockout. At the *OsGS3*-sgR01 target site, FrCas9 generated 69.6% (16 out of 23 T0 lines) indel frequency (Figure 2b). Among the T0 lines, 37.8% (8 out of 23) were biallelically edited, such as lines *osgs3-m1*, *osgs3-m2*, and *osgs3-m3* (Figure 2g). As expected, *OsGS3* biallelic mutants all showed increased seed length but not seed width or seed thickness (Figure 2h–k). Altogether, these data strongly support that CRISPR-FrCas9 confers efficient genome editing in stable T0 rice plants.

TREX2-FrCas9 generates larger deletions without compromising editing efficiency

Our data showed that FrCas9 predominantly generates 1 bp insertions (Figure 1) and 1 bp deletions (Figure 1j). While such

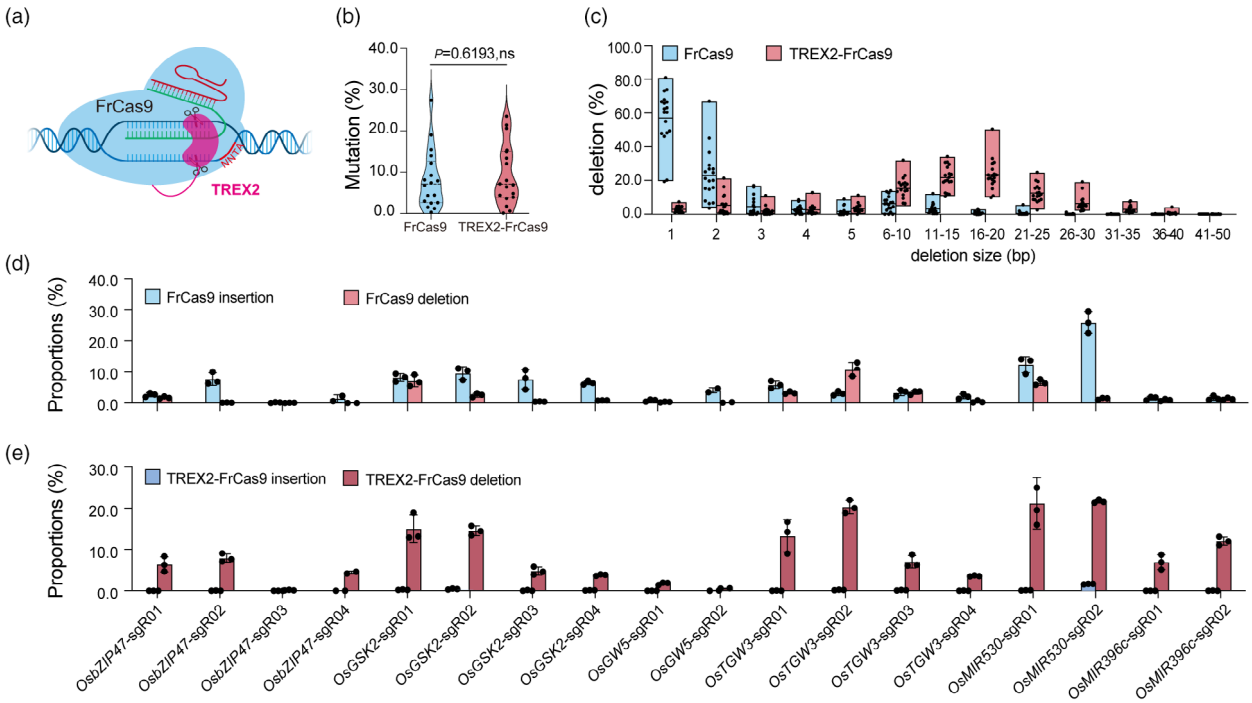


Figure 3 TREX2-FrCas9 generates larger deletions without compromising editing efficiency. (a) Schematic of the TREX2-FrCas9 system, which causes larger cuts through TREX2 exonuclease. (b) Comparing the mutation rates of the FrCas9 and TREX2-FrCas9 systems in rice protoplasts at 18 endogenous sites. Each dot represents a biological replicate of an independent experiment. Each target contains three biological replicates (n = 3). Data are presented as a violin plot, with line at the median. Statistical significance is indicated by asterisks. ns, $P > 0.05$, Student's t test. (c) Comparing the deletion sizes of the FrCas9 system in rice protoplasts at 18 endogenous sites. Each dot represents a biological replicate of an independent experiment. Each target contains three biological replicates (n = 3). Data are presented as floating bars (min to max), with a line at the mean. (d) The insertion and deletion proportions by FrCas9 at 18 rice loci. Each dot represents a biological replicate. Each assay contains three independent experiments (n = 3). Data are presented as mean values ± SD. (e) The insertion and deletion proportions by TREX2-FrCas9 at 18 rice loci. Each dot represents a biological replicate. Each assay contains three independent experiments (n = 3). Data are presented as mean values ± SD.

small indels can generate knockout in protein-coding genes, they would fall short in rendering strong effects of editing non-coding genes (e.g. microRNAs) or non-coding regions (e.g. cis-regulatory elements). In order to augment the editing profiles by increasing the deletion sizes, we tested a strategy in which an exonuclease domain is fused to FrCas9. Specifically, we fused the N-terminus of FrCas9 with Three Prime Repair Exonuclease 2 (TREX2), which is a 3′-to-5′ exonuclease (Cermak *et al*., 2017; Certo *et al*., 2012), and a 3XGGGGS linker was used (Figure 3a, Figure S5a). TREX2-FrCas9 was compared to the wild-type FrCas9 at 18 endogenous target sites in rice protoplasts. The data showed that TREX2-FrCas9 and FrCas9 had comparable editing efficiency at 13 out of 18 target sites (Figure S5b). At the remaining 5 target sites, the editing efficiency of TREX2-FrCas9 is higher than that of FrCas9 at 3 target sites (Figure S5b). Overall, TREX2-FrCas9 showed comparable editing efficiency to FrCas9 (Figure 3b). By comparing the deletion profiles at 18 rice endogenous sites, we found FrCas9 predominantly generated 1 bp deletions, followed by 2 bp deletions (Figure 3c). By contrast, TREX2-FrCas9 generated much larger deletions of 6–10 bp or longer, and the deletion sizes peaked at 16–20 bp (Figure 3c). TREX2-FrCas9 barely generated 1 bp deletions, but frequently generated deletions of over 16 bp (Figure 3c). The trend of generating larger deletions by TREX2-FrCas9 was evident when detailed analysis was done at all 18 sites (Figure S5c). Furthermore, while FrCas9 generated insertions and deletions at variable levels across different target sites (Figure 3d), there were barely any detectable insertions in TREX2-FrCas9 edited samples (Figure 3e). These data suggest that TREX2-FrCas9 drastically changes the editing profiles of the wild-type FrCas9 system, making TREX2-FrCas9 somewhat like CRISPR-Cas12a and Cas12b systems, which also predominantly generate larger deletions. Impressively, TREX2-FrCas9 achieved this capability without an overall compromise in editing efficiency.

TREX2-FrCas9-mediated deletion of *OsMIR156j* in rice

MicroRNAs (miRNAs) play a crucial role in post-transcriptional regulation. Previously, we established CRISPR-Cas9 as a useful tool for the genetic study of miRNA genes in plants. That study showed that 1 bp indels, as would be generated by SpCas9, were often not sufficient to destroy miRNA production at the target loci. As TREX2-FrCas9 can generate larger deletions, we reasoned that TREX2-FrCas9 would be an efficient tool for knocking out miRNA genes in plants. To demonstrate this, we developed a multiplexed TREX2-FrCas9 expression system where the sgRNAs are processed by tRNAs (Xie *et al*., 2015) (Figure 4a). Four sgRNAs were co-expressed with this system to target the *OsMIR156j* gene at different positions (Figure 4b). According to miRbase (http://www.mirbase.org/), we obtained the second structure of *pre-OsMIR156j*, which includes mature *OsMIR156j-5p* and *OsMIR156j-3p*. The four sgRNAs included two sgRNA pairs based on the palindromic PAM, with protospacers flanking both sides of the PAM site in each case (Figure 4b). We expected the multiplexed TREX2-FrCas9 construct to generate large deletions either by the action of individual sgRNAs or by the

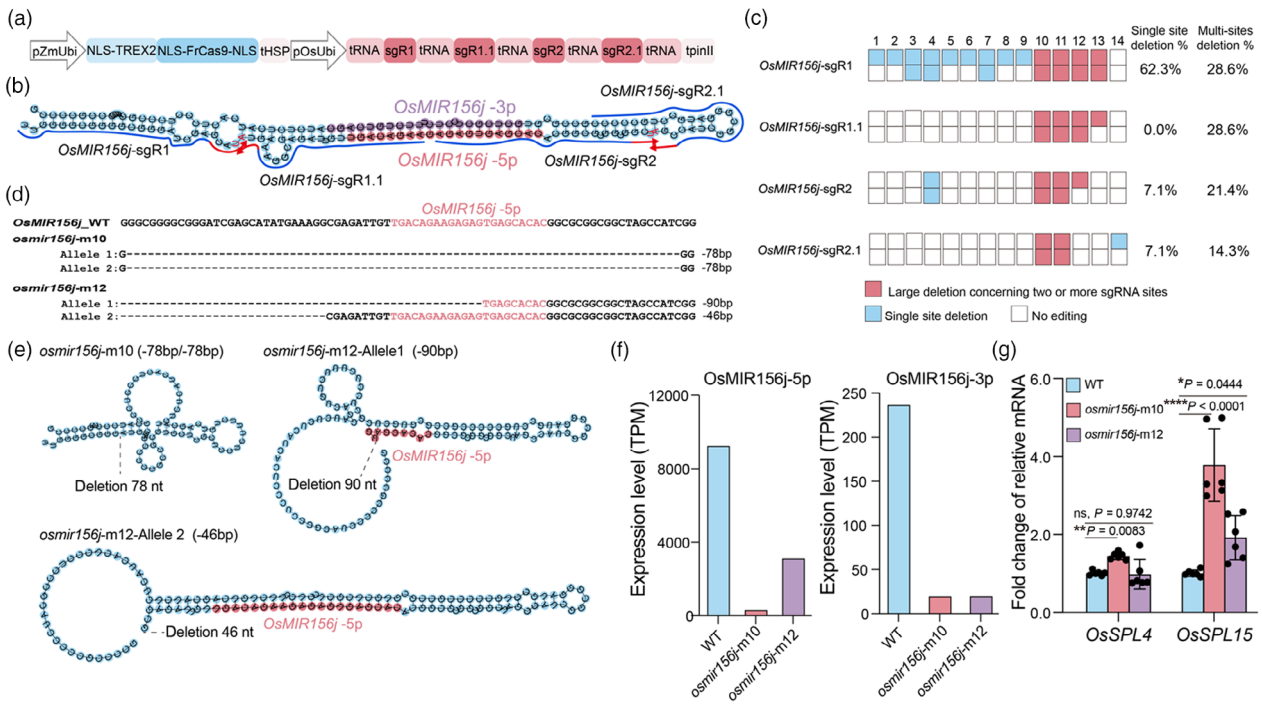
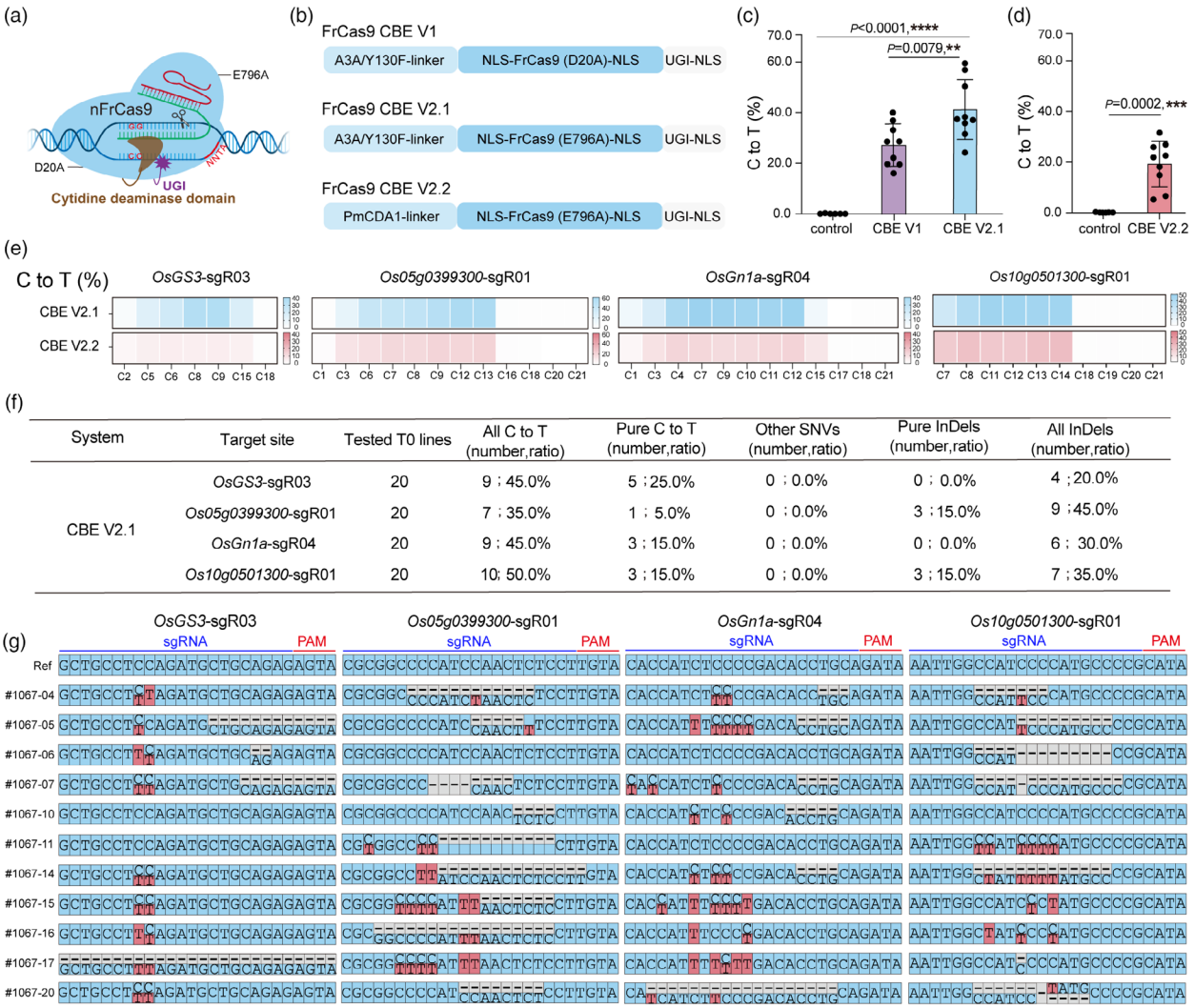


Figure 4 TREX2-FrCas9-mediated deletion of OsMIR156j in rice. (a) Diagrams of the *OsMIR156j* knockout vector based on TREX2-FrCas9. (b) Stem-loop structures of *OsMIR156j*. The mature miRNAs are highlighted in red and purple. The *OsMIR156j*-sgR1, *OsMIR156j*-sgR1.1, *OsMIR156j*-sgR2, and *OsMIR156j*-sgR2.2 were designed to target the *pre-OsMIR56j* stem-loop sites. (c) A table showed the genome editing efficiency of TREX2-FrCas9 in stable rice lines at *OsmiR156j* target sites. (d) Genotypes of two mutants derived from TREX2-FrCas9-mediated editing in mature *OsMIR156j*. (e) The secondary structure of pre-*OsMIR156j* mutants. (f) Transcript levels of *OsMIR156j-5p* and *OsMIR156j-3p* in the WT and the *OsMIR156j* mutants according to the small RNA seq experiment. (g) The expression profiles of *OsSPL4* and *OsSPL15*, respectively, based on qPCR data. Statistical significance is indicated by asterisks. ns, $P > 0.05$; *$P < 0.05$; **$P < 0.01$; ****$P < 0.0001$, one-way ANOVA, Dunnett's test.

combinational editing effects of more than one sgRNA. Fourteen stable transgenic rice lines showed editing at least at one target site (Figure 4c). Among the four sgRNAs, *OsMIR156j*-sgR1 is the most efficient, leading to genome edits at 13 out of 14 lines by this sgRNA (Figure 4c). Lines *osmir156j*-m10 and *osmir156j*-m11 showed biallelic deletions of all four adjacent target sites (Figure 4c, Figure S6). We selected *osmir156j*-m10 and *osmir156j*-m12 mutants for further study (Figure 4d, Figure S6). We used the web-based prediction tool (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) to predict the secondary structure of WT miRNA and miRNAs generated in the two mutants. The results showed significant changes in the secondary structure of *OsMIR156j* mutants (Figure 4e), and the *osmir156j*-m10 allele and *osmir156j*-m12-Allele 1 are both likely to be null alleles.

To get a better picture of the transcriptional regulation landscape shaped by the *OsMIR156j* knockout, we conducted RNA sequencing (RNA-seq) and small RNA-seq experiments with *osmir156j*-m10, *osmir156j*-m12, and the WT control (Figure 4d). For RNA-seq, we obtained over 45 million clean reads for each sample, with the mapped read coverage over 96% (Table S1), indicating high coverage. For small RNA-seq, we obtained over 4 million reads per sample. More than 700 small RNAs were identified in each sample, including >300 known miRNAs and ~420 novel miRNAs (Table S2). Interestingly, the expression level of pri-miR156j was down-regulated in the *osmir156j*-m10 mutant (Figure S7). Based on the miRNA prediction, both *OsMIR156j-5p* and *OsMIR156j-3p* will be produced (Figure 4b). Indeed, small RNA-seq detected *OsMIR156j-5p* (with an expression level of 9242 transcripts per million (TPM)) and *OsMIR156j-3p* (with an expression level of 235 TPM) in the WT plant (Figure 4f). Both miRNA forms were nearly undetectable in the *osmir156j*-m10 mutant (Figure 4f). Interestingly, in the *osmir156j*-m12 mutant, although *OsMIR156j-3p* was largely undetectable, *OsMIR156j-5p* showed significant residual expression (with an expression level of 3117 TPM) (Figure 4f). This is consistent with the prediction model that suggests *osmir156j*-m12-allele 2 may allow for the expression of intact *OsMIR156j-5p* (Figure 4e). In short, we confirmed genetic knockout and knock-down of the target miRNAs in both mutants.

Generating *OsMIR156j-5/3p* knockout and knockdown mutants provided us a good opportunity to validate their target genes. We made predictions on commonly used miRNA target gene prediction websites (https://www.zhaolab.org/psRNATarget/) and obtained a total of 181 potential target genes. In the RNA-seq results, 126 genes were detected, with 75 of them showing upregulation. Among these upregulated genes, several *OsSPL* genes are included (Figure S8). We further selected *OsSPL4* and *OsSPL15* validated by qRT-PCR according to predicted outcomes and expression levels (Figure 4g). These data suggest *OsMIR156j-5p*, which was not fully knocked out in *osmir156j*-m12, plays a significant role in repressing the expression of *OsSPL4* and *OsSPL15*. Together, these results support *OsMIR156j-OsSPL* regulation modules in rice as previously proposed (Jiao *et al.*, 2010; Shao *et al.*, 2019; Tseng *et al.*, 2023).



(f)

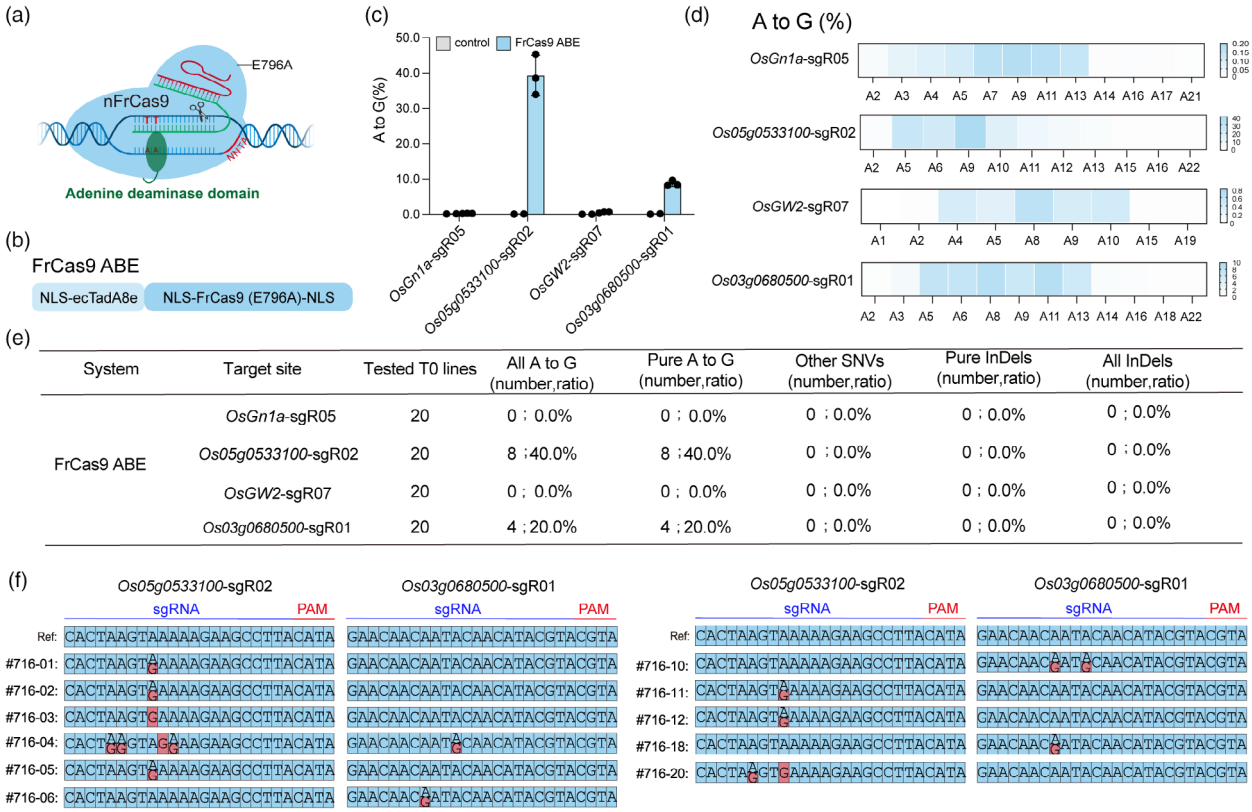
System	Target site	Tested T0 lines	All C to T (number,ratio)	Pure C to T (number,ratio)	Other SNVs (number,ratio)	Pure InDels (number,ratio)	All InDels (number,ratio)
CBE V2.1	*OsGS3*-sgR03	20	9 ; 45.0%	5 ;25.0%	0 ; 0.0%	0 ; 0.0%	4 ;20.0%
	Os05g0399300-sgR01	20	7 ; 35.0%	1 ; 5.0%	0 ; 0.0%	3 ;15.0%	9 ;45.0%
	OsGn1a-sgR04	20	9 ; 45.0%	3 ; 15.0%	0 ; 0.0%	0 ; 0.0%	6 ; 30.0%
	Os10g0501300-sgR01	20	10; 50.0%	3 ; 15.0%	0 ; 0.0%	3 ; 15.0%	7 ;35.0%

Figure 5 Efficient FrCas9-based cytosine base editors. (a) Schematic of the FrCas9-based CBE system, which causes C-G to T-A. (b) Schematic representation of three FrCas9-based CBE systems in rice. (c) Assessment of FrCas9 CBE V1 and FrCas9 CBE V2.1 editors in rice protoplasts at four CG-rich sites with next-generation sequencing (NGS) of PCR amplicons. Each target contains at least two biological replicates. Data are presented as mean values ± SD. Statistical significance is indicated by asterisks. ns, $P > 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$, one-way ANOVA, Dunnett's test. (d) Assessment of FrCas9 CBE V2.2 in rice protoplasts at four CG-rich sites with NGS of PCR amplicons. Each target contains at least two biological replicates. Data are presented as mean values ± SD. Statistical significance is indicated by asterisks. ns, $P > 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$, Student's *t* test. (e) Editing windows heatmap of cytosine base conversions induced by FrCas9 CBE V2.1 and FrCas9 CBE V2.2 editors at four CG-rich sites. The substitution frequency of each base was calculated as the percentage of edited reads within the total reads. (f) Base editing frequency of FrCas9 CBE V02.1 in stable rice lines. (g) Genotypes of all T0 lines containing genome editing events in (f).

FrCas9-based cytosine base editors confer efficient C-to-T base editing

CRISPR-Cas-derived base editors are powerful tools for rapid, efficient, and accurate genetic perturbation and improvements in plants. Base editors have been made from type II CRISPR-Cas9 or type V CRISPR-Cas12a systems to achieve targeted, precise nucleotide substitutions without double-strand breaks (DSBs) (Cheng *et al.*, 2023; Molla *et al.*, 2021). To test whether base editors can be developed from FrCas9, we compared two FrCas9 nickase versions: FrCas9 (D20A), which corresponds to D10 of the RuvCI domain in SpCas9 and FrCas9 (E796A) which is used in human cells (Figures 1a and 5a) (Cui *et al.*, 2022). We adopted a BE3 configuration by fusing the N-terminus of these nickases with cytidine deaminase A3A/Y130F (Figure 5b), which was shown to be very efficient in rice (Ren *et al.*, 2021b), tomato (Randall *et al.*, 2021), and poplar (Li *et al.*, 2021). These two CBE versions (CBE V1 and V2.1) of FrCas9 were constructed for testing in rice protoplasts at four target sites, each with a 22 bp spacer in the optimal sgRNA V02 scaffold. We assessed base editing efficiency by amplicon-based NGS. The data showed that CBE V2.1 generated up to 59.5% C-to-T conversion, which is higher than the 40.0% C-to-T conversion rate achieved by CBE V1 (Figure 5c). Interestingly, CBE V2.1 was based on FrCas9 (E796A), not FrCas9 (D20A), whose equivalent SpCas9 (D10A) is widely used for developing efficient base editors.

Based on the FrCas9-E796A nickase, we further replaced A3A/Y130F with PmCDA1 (Nishida *et al.*, 2016), which is an



(e)

System	Target site	Tested T0 lines	All A to G (number,ratio)	Pure A to G (number,ratio)	Other SNVs (number,ratio)	Pure InDels (number,ratio)	All InDels (number,ratio)
FrCas9 ABE	*OsGn1a*-sgR05	20	0 ; 0.0%	0 ; 0.0%	0 ; 0.0%	0 ; 0.0%	0 ; 0.0%
	Os05g0533100-sgR02	20	8 ;40.0%	8 ;40.0%	0 ; 0.0%	0 ; 0.0%	0 ; 0.0%
	OsGW2-sgR07	20	0 ; 0.0%	0 ; 0.0%	0 ; 0.0%	0 ; 0.0%	0 ; 0.0%
	Os03g0680500-sgR01	20	4 ;20.0%	4 ;20.0%	0 ; 0.0%	0 ; 0.0%	0 ; 0.0%

Figure 6 Efficient FrCas9-based adenine base editors. (a) Schematic of the FrCas9-based ABE system, which causes A-T to G-C. (b) Schematic representation of the FrCas9-based ABE system in rice. (c) Assessment of FrCas9 ABE editors in rice protoplasts at four AT-rich sites. Each dot represents a biological replicate. Each target contains at least two biological replicates. Data are presented as mean values ± SD. (d) Editing windows heatmap of adenine base conversions induced by FrCas9 ABE editors at four AT-rich sites. The substitution frequency of each base was calculated as the percentage of edited reads within the total reads. (e) Base editing frequency of FrCas9 ABE in stable rice lines. (f) Genotypes of all T0 lines containing genome editing events in (e).

efficient cytidine deaminase with undetectable off-target effects in plants (Ren *et al.*, 2021b). The resulting CBE V2.2 version was also tested at the same four target sites. The data showed that CBE V2.2 generated up to a 31.5% C-to-T conversion rate in rice protoplasts (Figure 5d). Further analysis showed that both CBE V2.1 and CBE V2.2 showed very minimal levels of indel byproducts in rice protoplasts (Figure S9). CBE V2.1 and CBE V2.2 both showed wide base editing windows spanning from the 3rd to the 15th nucleotide in the 5′ to 3′ direction within the protospacer (Figure 5e).

To see whether the efficient CBE V2.1 can produce base-edited plants at the four target sites, we carried out stable rice transformation and analysis. Genotyping of the T0 lines revealed robust cytosine base editing by CBE V2.1, generating ~50% editing efficiency at all four target sites (Figure 5f). Further analysis of the edits showed the events of pure C-to-T editing as well as indel mutations in T0 lines (Figure 5f). It is well known that indels can be generated by CBEs (Komor *et al.*, 2016), and indel formation could be further inhibited by the expression of more UGI, as we previously showed (Ren *et al.*, 2021b). Nevertheless, monoallelic and biallelic pure C-to-T base editing lines can be identified within the T0 population at all four target sites (Figure 5g). Therefore, FrCas9-based CBE V2.1 is an efficient cytosine base editor for genome editing in plants.

Development of an adenine base editor based on FrCas9 and TadA-8e

Based on the information in developing FrCas9 CBE systems, we reasoned that an efficient adenine base editor (ABE) could be made by coupling the high-processing ecTadA-8e (Richter *et al.*, 2020) and the FrCas9 (E796A) nickase, resulting in FrCas9 ABE (Figure 6a,b). We tested the base editing efficiency of FrCas9 ABE with a 22-bp spacer and sgRNA V02 at four endogenous sites in the rice genome. Base editing efficiency in rice protoplasts was assessed by amplicon-based NGS. The data showed that the FrCas9 ABE could achieve up to 39.3% A-to-G conversion (Figure 6c). However, base editing efficiency at two target sites, *OsGn1a*-sgR05 and *OsGW2*-sgR07, was extremely low (<1%) (Figure 6c). FrCas9 ABE showed very minimal levels of indel byproducts (Figure S10). It showed a base editing window spanning from the 4th to the 13th nucleotide in the 5′ to 3′ direction within the protospacers (Figure 6d). Further testing of the FrCas9 ABE system in stable transgenic rice plants showed editing efficiency consistent with the data obtained in rice protoplasts. At the *Os05g0533100*-sgR02 and *Os03g0680500*-sgR01 sites, A-to-G editing efficiencies of 40.0% and 20.0% were achieved (Figure 6e). At the two target sites with very low base editing efficiencies in protoplasts, no editing events were recovered in the T0 lines (Figure 6e). These data suggest that the

prosperity of FrCas9 ABE in stable rice plants can be predicted using the data from the rice protoplasts. Impressively, in those lines with A-to-G base editing events (either monoallelic or biallelic), no indel mutants were found (Figure 6f), suggesting high editing purification of FrCas9 ABE.

Off-target assessment of the CRISPR-FrCas9 systems

For any new CRISPR-Cas genome editing system, it is important to assess its potential off-target effects. In base editing, the most concerning off-target efforts result from guide RNA-independent off-target effects due to non-specific deaminase activities at relatively random sites in the genomes. The cytidine deaminases A3A/Y130F and PmCDA1 were previously assessed by WGS in rice, and the off-target effects are minor for A3A/Y130F and none for PmCDA1 (Ren *et al.*, 2021b). Furthermore, analysis in tomato showed undetectable off-target effects by A3A/Y130F at both DNA and RNA levels (Randall *et al.*, 2021). The adenine deaminase ecTadA-8e was shown to generate off-target mutations at DNA and RNA levels in rice but generated undetectable off-target effects in tomato (Li *et al.*, 2022; Sretenovic *et al.*, 2023a; Wu *et al.*, 2022b). Since the off-targeted effects of the deaminases used in the development of FrCas9 base editors have been previously investigated, we focused on our off-target analysis of FrCas9 and TREX2-FrCas9 nucleases, as it would allow us to assess (i) the off-target effects of FrCas9 and (ii) the off-target effects of TREX2.

We employed a whole genome sequencing (WGS) pipeline to thoroughly evaluate the genome-wide off-target effects of the FrCas9 and TREX2-FrCas9 editors (Figure 7a). A total of 18 plants were selected for WGS, including 10 lines edited by FrCas9 editors and 4 lines edited by TREX2-FrCas9 (Figure 7b, Table S3). Additionally, two wild-type (WT) plants and two Agrobacterium transformation (Agro) control plants were included in the analysis (Figure 7a). Sequencing reads of all plants were mapped to the rice genome (Table S4). The number of indels, ranging from approximately 500 to 1000, observed in all plants, whether with or without sgRNAs, exhibited similarity to those found in control plants (Figure 7c). This finding implies that these indels are attributed to somaclonal variation arising from tissue culture. In Agro-only and FrCas9 plants, comparable numbers of single nucleotide variations (SNVs) were identified, averaging around 250 (Figure 7d). Interestingly, plants that expressed TREX2-FrCas9 carried a substantially higher number of SNVs, with an average of about 700 SNVs per plant (Figure 7d). To further investigate this off-target effect, we conducted an analysis of all six potential nucleotide substitutions detected as SNVs (Figure 7e). In comparison to Agro-only control plants, the plants expressing FrCas9 exhibited no significant change in the quantity of these six potential nucleotide substitutions, suggesting an undetectable off-target effect for FrCas9. However, plants expressing TREX2-FrCas9 showed a significant increase in the number of C:G>T:A and A:T>T:A mutations, rising from 90 to nearly 200 and from 30 to almost 300, respectively. This indicates that TREX2-FrCas9 results in a significant increase in nucleotide substitution mutations, preferably causing C:G>T:A and A:T>T:A off-target mutations (Figure 7e,f). The observed SNVs were distributed randomly across all 12 chromosomes of the rice genome (Figure 7g).

We also utilized the CRISPR RGEN Tools (http://www.rgenome.net/cas-offinder/) to predict potential off-target sites that are dependent on the gRNA. We analysed the number of off-target sites with a mismatch <5 bp in each targeted site and analysed genotype at some of these sites in FrCas9 and TREX2-FrCas9 edited plants (Figure S12a). The results showed that no mutations occurred in these gRNA-dependent off-target sites (Figure 7h), even at the putative off-target *OsGn1a*-sgR01-OF site, which shared the same protospacer as the on-target site but with an altered 5′-CTCA-3′ PAM. Overall, both FrCas9 and TREX2-FrCas9 exhibit high specificity across the entire genome (Figure 7h). Considering the importance of PAM sequences, we further investigated potential gRNA+PAM-dependent off-target sites in more plants. By examining the predicted off-target sites of *OsEPFL9*-sgR01 and *OsGS3*-sgR01 by FrCas9 and those of the four target sites of *OsMIR156j* by TREX2-FrCa9, no off-target events were identified at any of the predicted off-target sites based on Sanger sequencing (Figure S11b,c).

Taken together, we did not find editing at any of the predicted off-target sites in the plants edited by FrCas9 and TREX2-FrCas9, suggesting highly specific gRNA-dependent genome editing by these systems. Undetectable gRNA-independent off-target effects in the rice genome suggest FrCas9 is a highly specific nuclease. The fusion of TREX2 to FrCas9 can cause gRNA-independent off-target mutations, suggesting an off-target effect linked to TREX2 expression in plants.

Discussion

CRISPR-Cas9, from the type II CRISPR-Cas bacterial adaptive immune system, has been engineered as an efficient genome editing platform for a wide range of organisms (Sander and Joung, 2014). CRISPR-Cas9 and its derived base editors and prime editors have revolutionized plant reverse genetics and crop breeding (Gao, 2021; Molla *et al.*, 2021; Zhou *et al.*, 2019). Innovative applications have gone beyond single gene knockout, as demonstrated in *de novo* crop demonetization based on multi-gene editing (Curtin *et al.*, 2022; Yu *et al.*, 2021) and engineering quantitative trait variation by multiplexed promoter editing (Rodriguez-Leal *et al.*, 2017; Tang and Zhang, 2023). The widely used *Streptococcus pyogenes Cas9* (SpCas9) primarily recognizes the 5′-NGG-3′ PAM sequence (Jinek *et al.*, 2012), which limits its targeting ranges. Therefore, it is of great interest and importance to develop robust CRISPR genome editing systems that recognize alternative PAMs.

Two potential approaches have been used to identify CRISPR-Cas9 systems that confer genome editing at alternative PAM sites. In the first approach, more Cas9 orthologs have been discovered and demonstrated, such as SaCas9, St1Cas9, and St3Cas9 from the Type II-A CRISPR family (Hu *et al.*, 2020; Qin *et al.*, 2019; Ran *et al.*, 2015; Xu *et al.*, 2020) and Nm1Cas9, Nm2Cas9, GeCas9, CcCas9, CdCas9, and CjCas9 from the Type II-C family (Chylinski *et al.*, 2014; Fedorova *et al.*, 2020; Hirano *et al.*, 2019; Hou *et al.*, 2013; Xu *et al.*, 2022; Zhu *et al.*, 2019). In the second approach, SpCas9 variants have been engineered, such as SpCas9-VQR (Hu *et al.*, 2016), SpCas9-EQR (Hu *et al.*, 2016), SpCas9-VRER (Qin *et al.*, 2019), Cas9-NG (Zhong *et al.*, 2019), and SpRY (Ren *et al.*, 2021c; Sretenovic *et al.*, 2023b). Among these Cas9 nucleases, Cas9-NG and SpRY are promising due to their simple PAM requirements. However, such relaxed PAM requirements seem to be achieved at the cost of editing robustness and efficiency, as indicated by many recent studies in plants (Ren *et al.*, 2019, 2021a,b; Sretenovic *et al.*, 2023b; Xu *et al.*, 2021; Zhong *et al.*, 2019). It is thus important to search for additional Cas9 nucleases that confer high-efficiency genome editing with simple PAM requirements.

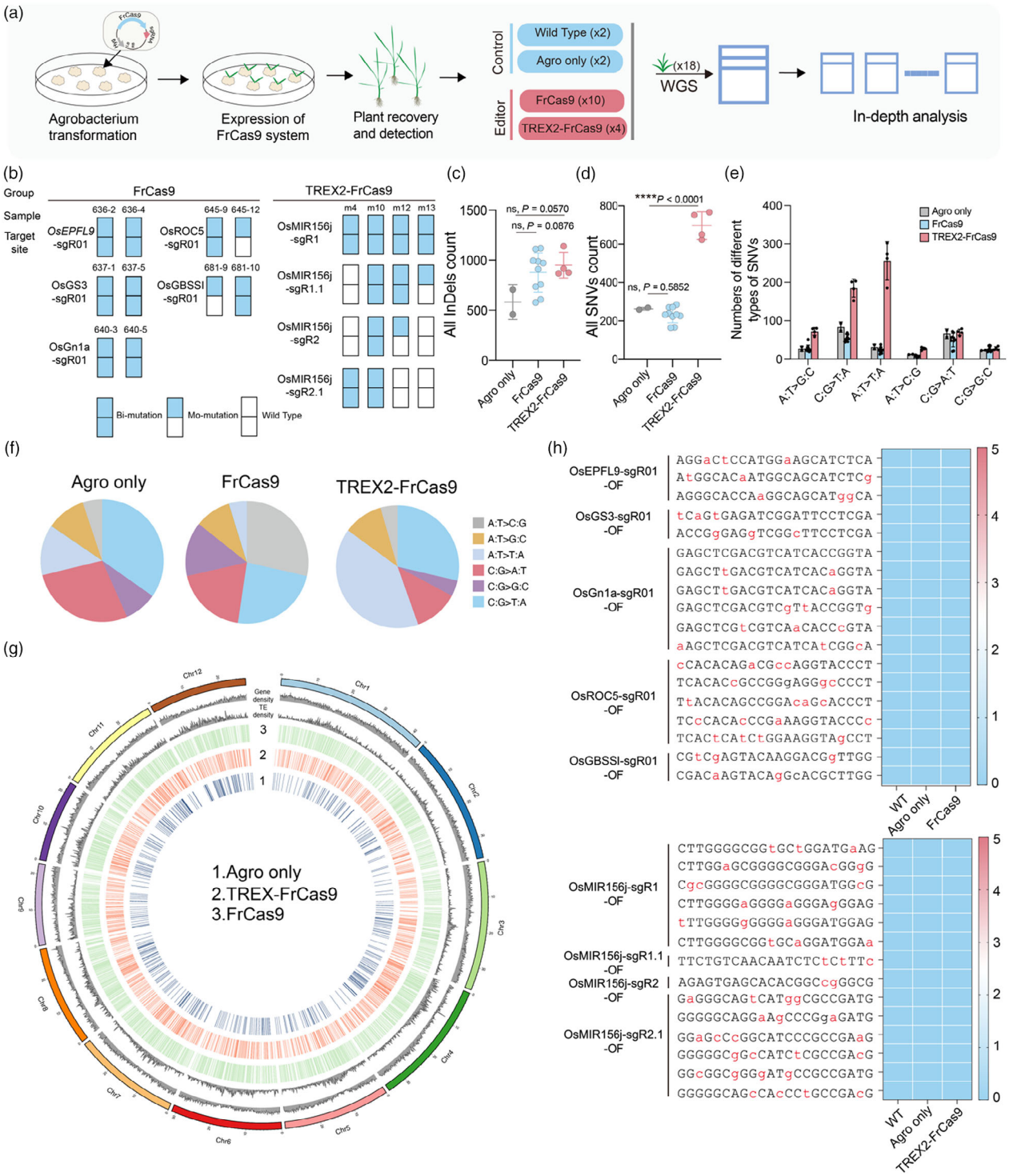


Figure 7 Genome-wide off-target analysis of FrCas9 editors. (a) The pipeline for whole genome sequencing (WGS)-based off-target analysis. (b) Summary of the rice lines used for WGS. (c) Number of indels mutations in all sequenced samples. Each dot represents a biological replicate. Each target contains two biological replicates. (d) Number of SNV mutations in all sequenced samples. Each dot represents a biological replicate. Each target contains at least two biological replicates. Data are presented as mean values ± SD. Statistical significance is indicated by asterisks. ns, $P > 0.05$; ****$P < 0.0001$, one-way ANOVA, Dunnett's test. (e) Frequency analysis of each base change type among each sample type identified by WGS. (f) Proportion of different mutation types in SNVs (Pie Chart). (g) Genome-wide distribution of SNVs in all sequenced samples. (h) Mutation analysis of predicted gRNA-dependent off-target sites based on WGS data; blue indicates no off-target mutation.

In addition to the aforementioned type II CRISPR-Cas proteins, many type V CRISPR-Cas proteins have been explored for genome editing, including Cas12a (Cheng *et al.*, 2023; Gaillochet *et al.*, 2023; Tang *et al.*, 2017; Zetsche *et al.*, 2015, 2020; Zhang *et al.*, 2021), Cas12b (Ming *et al.*, 2020), Cas12c (Harrington *et al.*, 2020), Cas12d (Harrington *et al.*, 2020), Cas12f (Kong

et al., 2023), Cas12i (Lv *et al*., 2023; Zhang *et al*., 2023), Cas12m (Wu *et al*., 2022a), and Cas12j (Liu *et al*., 2022). These Cas12 nucleases generally recognize 5′-NTTN-3′ PAMs and generate larger deletions due to sticky ends produced after DNA cleavage. These two characteristics make CRISPR-Cas12 systems suitable tools to target A-T-rich regions, such as promoters in plants. This has been demonstrated in our research promoter editing study that utilized LbCas12a and an efficient model for target site prediction (Zhou *et al*., 2023). Moreover, there is significant interest in developing CRISPR-Cas9 genome engineering systems capable of targeting A-T-rich PAM sites. Recently, we developed a CRISPR-LrCas9 for genome engineering in plants with a 5′-NGAAA-3′ PAM requirement (Zhong *et al*., 2023). LrCas9 has a complementary PAM to those of Cas12a proteins, making it a potentially useful tool for targeting A-T-rich promoters. However, unlike Cas12a nucleases, LrCas9 only generated small indels, rendering it a less efficient tool for promoter editing and reverse genetics of non-coding genes. Furthermore, the PAM of LrCas9 is still too complicated when compared to SpCas9.

In this study, we were motivated to explore FrCas9 from *Faecalibaculum rodentium* as it has a simple 5′-NNTA-3′ PAM (Cui *et al*., 2022). The simple and palindromic nature of the 5′-NNTA-3′ PAM makes FrCas9 a useful tool for genome editing in plants. Indeed, our *in silico* analysis showed that FrCas9 can target more sites than SpCas9 in the rice genome (Figure 1d), which is probably true for other plants as well. To make FrCas9 a more useful genome editing tool, we generated TREX2-FrCas9. Thanks to the exonuclease activity of TREX2, the TREX2-FrCas9 fusion predominantly produces larger deletions (Figure 3c), just like the widely used CRISPR-Cas12a systems (Tang *et al*., 2017; Zhang *et al*., 2021; Zhong *et al*., 2018). The deletion generated by Cas12a is concentrated between 6 bp and 13 bp, with a peak of 8–9 bp (Tang *et al*., 2017). In contrast, the deletion generated by TREX2-FrCas9 is mainly >11 bp, with a peak of 16–20 bp deletions. Hence, TREX2-FrCas9 could be a superior tool for editing non-coding cis elements and non-coding genes. Due to the simple AT-rich PAM requirement and larger deletion sizes, TREX2-FrCas9 is potentially more advantageous to LbCas12a for promoter editing (Zhou *et al*., 2023).

To demonstrate the novel use of TREX2-FrCas9, we focused on a case study of editing a complex *OsMIR156j* locus. By multiplexing four sgRNAs, we efficiently edited this miRNA locus. Rice mutants of *OsMIR156j* were generated. Further RNA-seq and small RNA-seq experiments confirmed the loss of function of this miRNA locus. In these *OsMIR156j* rice mutants, changes in the transcription levels of several *OsSPL* genes, such as *OsSPL4* and *OsSPL15*, were observed, aligning with the previously proposed models, thereby indicating their potential regulation by *OsMIR156j* (Jiao *et al*., 2010; Shao *et al*., 2019; Tseng *et al*., 2023). Thus, our study demonstrated TREX2-FrCas9 as an efficient tool for genetic knockout of miRNA genes in plants.

Because of the palindromic structure of the TA sequence, FrCas9 was investigated through the co-expression of two sgRNAs with a shared PAM in our *OsMIR156j* study. The use of dual sgRNA may result in the deletion of multiple fragments. The four sgRNAs targeting *OsMIR156j* were two pairs that were designed based on the palindromic nature of the 5′-NNTA-3′ PAM of FrCas9 (Figure 4b). Although these four sgRNAs were co-expressed by the tRNA-based multiplexed system (Figure 4a), analysis of the editing outcomes in the T0 rice lines showed OsMIR156j-sgR1 is capable of inducing mutations in *OsMIR156j*, demonstrating its efficiency (Figure 4c). This indicates that when designing two sgRNAs to flank a palindromic 5′-NNTA-3′ PAM site for creating knock-out mutants, it is advisable to assess the activity of each sgRNA through a protoplast transient expression assay. Subsequently, utilizing both effective sgRNAs can enhance the overall editing efficiency. The one with better binding affinity and efficiency would win. After all, FrCas9 needs to bind to the overlapping PAM sites, and maybe the steric hindrance will prevent simultaneous editing by two closely positioned sgRNAs. Because FrCas9 editing efficiency, as with other Cas9 systems, is largely dependent on the protospacer sequences (Figure 1h), it is probably still a good idea to design paired sgRNAs utilizing the same palindromic PAM site, knowing the editing outcomes will be dictated by the better sgRNA. Furthermore, given that CRISPR-Cas-based genome editing is affected by chromatin status and epigenomic marks (Li *et al*., 2013; Liu *et al*., 2019; Weiss *et al*., 2022), the palindromic nature of FrCas9's PAM provides a unique opportunity to assess many pairs of sgRNAs with each targeting the same PAM site. In such an experiment, the only variable would be the composition of protospacers. This would allow for massive parallel analysis of many sgRNA pairs to figure out all the winners and losers in paired analysis. We envision that obtaining such big data about genome editing with paired sgRNAs would enable machine learning to help design more efficient sgRNAs, something worth pursuing in the future.

We also developed CBEs- and ABEs-based FrCas9 for achieving C-to-T base editing and A-to-G base editing in plants. Interestingly, FrCas9 (E796A) appeared to be a better nickase than FrCas9 (D20A) for engineering efficient base editors (Figure 5c). This is a bit surprising as the counterpart of the D20A mutation in SpCas9, D10A, is nearly exclusively used to develop all base editors from SpCas9 (Molla *et al*., 2021). However, our data are consistent with the findings about FrCas9 nickases in human cells (Cui *et al*., 2022). We found FrCas9 CBE based on A3A/Y130F is more efficient than that based on PmCDA1, which is also consistent with similar results obtained with SpCas9 (Ren *et al*., 2021b). Interestingly, FrCas9 CBE V2.2 based on PmCDA1 showed a wide base editing window, similar to that of FrCas9 CBE V2.1 based on A3A/Y130F (Figure 5e). Previously, we showed that PmCDA1-based SpCas9 CBE has a narrow base editing window towards the 5′ end of the protospacer (Ren *et al*., 2021b). In that case, PmCDA1 was fused to the C-terminus of SpCas9-D10A. So, it will be interesting to test whether base editing windows can be shifted by fusing the cytidine deaminase to different ends of the Cas9 nickase. Analysis of FrCas9 CBE V2.1 showed frequent indel byproducts in T0 rice lines. It is possible that we can reduce indel formation by recruiting more copies of uracil glycosylase inhibitor (UGI), as we previously did with SpCas9 CBEs, where MS2-MCP interaction was used for UGI recruitment (Ren *et al*., 2021b). However, this would require further sgRNA scaffold engineering to identify the optimal insertion sites for MS2. By contrast, FrCas9 ABE did not generate detectable indels in both rice protoplasts and stable lines. Potential future improvements to FrCas9 ABE may focus on enhancing editing robustness.

Our comprehensive WGS analysis demonstrated that CRISPR-FrCas9 is a highly specific genome editing system in plants (Figure 7). It is worth noting that we observed a significant off-target effect induced by TREX2-FrCas9 in rice, resulting in a substantial number (~700) of whole-genome SNV mutations, with the majority being C:G>T:A and A:T>T:A substitutions (Figure 7d–g). Since such off-target mutations were not observed with FrCas9, those observed in TREX2-FrCas9-edited plants must

be caused by TREX2 expression. Given that dsDNA or chromosomal DNA is a natural substrate of TREX2 and TREX2 has a 3′ to 5′ exonuclease activity (Cheng *et al*., 2018), it is not too surprising that expression of TREX2 in plant cells can generate off-target mutations. However, it is intriguing to see that these off-target mutations are SNVs. The predominant C:G>T:A and A:T>T:A substitutions caused by TREX2 suggest its preference for causing C-to-T and A-to-T mutations when acting on DNA. It is plausible that TREX2 might interfere with DNA repair pathways. The mechanism behind these off-target SNV mutations is unclear and warrants further investigation. Due to this off-target effect, TREX2-FrCas9 edited plants had ~1600 mutations (indels and SNVs), which are more than ~1000 mutations observed in the Agro control plants and FrCas9 edited plants (Figure 7c,d). Considering the level of somaclonal variation (~1000 mutations in our experiments), the number of mutations observed in TREX2-FrCas9 plants is not alarmingly high. Previously, with WGS, we detected sgRNA-independent off-target mutations in rice plants edited by BE3-A3A/Y130F (Ren *et al*., 2021b). Yet, such off-target effects were undetectable in tomato plants by the same base editor (Randall *et al*., 2021). These observations suggest off-target effects may vary among plant species, which could be partly affected by the genome editor's expression levels. So, it is uncertain whether TREX2-FrCas9 would always lead to some genome-wide off-target effects in other plants. Regardless, this current level of off-target effects observed in rice should not prevent TREX2-FrCas9 from being used for a wide range of genome editing applications in plants, taking advantage of its large deletion profiles.

Conclusion

In this study, we developed an efficient FrCas9 system for plant genome editing. The system allows for targeted mutagenesis, C-to-T base editing, and A-to-G base editing. To augment the FrCas9 nuclease system, we generated TREX2-FrCas9, which produced large deletions without comprising editing activity. The usefulness of TREX2-FrCas9 was demonstrated in the editing and genetic analysis of a miRNA locus in rice. We expect these new FrCas9-based genome editing systems, with the unique feature of relying on a palindromic 5′-NNTA-3′ PAM, will greatly add to the plant genome editing toolbox and further aid genome editing-based crop breeding.

Experimental procedures

Plant materials

The *Japonica* cultivar Nipponbare was used in this study. The protoplast transformation materials are prepared by growing sterilized rice seeds in 1/2 MS culture medium for 10–12 days at 28 °C in the dark, resulting in rice seedlings suitable for protoplast transformation. For stable transformation, sterilized rice seeds are placed in callus induction medium (N6-D), as described previously (Zhong *et al*., 2019), under 32 °C with 24 h of light exposure. They are cultured for 7–10 days to obtain calli suitable for stable transformation of rice.

Statistics of the sgRNA number

A custom Python script (https://github.com/yuechaowu/Find_gRNA) was developed to calculate sgRNA numbers in the rice genome. For SpCas9, the target search parameter was set as 5′-nnnnnnnnnnnnnnnnnnnnNGG-3′ (20 nt protospacer + NGG PAM). For FrCas9, the target search parameter was set as 5′-nnnnnnnnnnnnnnnnnnnnNNTA-3′ (22 nt protospacer + NNTA PAM). Firstly, a search was done to identify targetable sgRNA sites throughout the entire rice genome. Then, the sgRNAs that target gene-coding regions were removed to reveal sgRNAs targeting the non-coding regions of the rice genome.

Construction of the vectors

The vectors were constructed based on the backbone of pTrans_210d (Addgene Plasmid #91109). The DNA sequence of the FrCas9 gene was rice codon-optimized and synthesized. Then, the FrCas9 DNA fragments were constructed into pTX770, which was constructed with pZmUbi1-SV40 NLS-*BsaI* site1-ccdb-*BsaI* site2-nucleoplasmin NLS-AtHSP by Golden Gate Assembly to generate pHY618 (MOD_A). To obtain the sgRNA expression cassette named pHY619 (MOD_B), the sgRNA are driven by an OsUbi1 promoter and further processed by tRNAs, resulting in pHY619: pOsUbi1-tRNA (Gly)-*BsaI* site3-ccdb-*BsaI* site 4-gRNA scaffold-tRNA-tpinII. We assembled the expression backbone pGEL846 (Addgene Plasmid #214236) by Golden Gate Assembly of pTrans_210d, MOD_A (pHY618), MOD_B (pHY619), and MOD_C (pMOD_C0000A) (Liu *et al*., 2022). For the FrCas9 singular editing vectors, the sgRNA was cloned into the expression backbone vector pGEL846 according to the previously established method (Zhou *et al*., 2023).

To prepare the backbone vector pHY847 (Addgene Plasmid #214237) of TREX2-FrCas9 editing systems, we first constructed the TREX2-FrCas9 expression vector pHY892. The DNA sequence of the TREX2-GGGGSx3_linker-FrCas9 gene was rice codon optimized and synthesized in one fragment (Mazur and Perrino, 1999). Then, the TREX2-GGGGS x3_linker-FrCas9 DNA fragments were constructed into pTX770 by Golden Gate Assembly to generate pHY892 (MOD_A). We assembled the expression backbone pGEL847 (Addgene Plasmid #214237) by Golden Gate Assembly of pTrans_210d, MOD_A (pHY892), MOD_B (pHY619), and MOD_C (pMOD_C0000A). TREX2-FrCas9 singular editing vectors were constructed by annealing four synthesized oligonucleotides flanked by *BsaI* restriction enzyme sites. The final T-DNA recombinant expression vectors were constructed with Golden Gate reactions.

To prepare the backbone vector pHY848 (Addgene Plasmid #214238) of FrCas9 CBE V2.1 editing systems, we first constructed the A3A/Y130-XTEN linker-nFrCas9 (E796A)-UGI expression vector pHY841 (Ren *et al*., 2021b). The A3A/Y130-XTEN linker-nFrCas9 (E796A)-UGI DNA fragments were constructed into pTX770 by Golden Gate Assembly to generate pHY841 (MOD_A). We assembled the expression backbone pGEL848 (Addgene Plasmid #214238) by Golden Gate Assembly of pTrans_210d, MOD_A (pHY841), MOD_B (pHY619), and MOD_C (pMOD_C0000A). The FrCas9 CBE V2.1 singular editing vectors were generated by annealing four synthesized oligonucleotides flanked by *BsaI* restriction enzyme sites. The final T-DNA recombinant expression vectors were constructed with Golden Gate reactions.

To prepare the backbone vector pHY849 (Addgene Plasmid #214239) of FrCas9 ABE editing systems, we first constructed the ecTadA8e-32aa linker-nFrCas9 (E796A) expression vector pHY710 (Cui *et al*., 2022). The ecTadA8e-32aa linker-nFrCas9 (E796A) DNA fragments were constructed into pTX770 by Golden Gate Assembly to generate pHY710 (MOD_A). We assembled the expression backbone pGEL849 (Addgene Plasmid #214239) by Golden Gate Assembly of pTrans_210d, MOD_A

(pHY710), MOD_B (pHY619), and MOD_C (pMOD_C0000A). The FrCas9 ABE singular editing vectors were generated by annealing four synthesized oligonucleotides flanked by *BsaI* restriction enzyme sites. The final T-DNA recombinant expression vectors were constructed with Golden Gate reactions.

Sanger sequencing was employed to confirm the integrity of all vectors. Table S5 provides a comprehensive list of the sgRNAs utilized in this study.

Rice protoplast transformation

The Japonica cultivar Nipponbare rice seedlings were grown on 1/2 MS solid medium for 10–12 days in the dark at 28 °C. The rice protoplast extraction and transformation methods were done by following our previously published protocols (Lowder *et al.*, 2015; Tang *et al.*, 2017). In summary, healthy leaves were finely sliced into 0.5–1.0 mm strips and immersed in an enzyme solution. Following a 30-min vacuum infiltration, the leaves were incubated in the dark at 25 °C under gentle agitation (60–80 rpm) for 6 h. The digestion mixture was then filtered through a 40-μm nylon mesh. After two washes with W5 washing buffer, the protoplasts were carefully examined and counted under a microscope. The final concentration of protoplasts was adjusted to 2×10^6 per millilitre. For the protoplast transformation, 30 μg of plasmid DNA in 30 μL (1 μg/μL, prepared using the Qiagen Midiprep kit) was gently mixed with 200 μL protoplasts and 230 μL of 40% PEG transformation buffer. After a 30-min incubation in the dark, the reactions were halted by adding 900 μL of W5 washing buffer. The protoplasts were centrifuged at a low speed and then transferred to a 12-well culture plate for further incubation in the dark at 32 °C for 48 h. In these experiments, the quality of each transformation was meticulously assessed through parallel experiments utilizing a fluorescence reporter. Only those control experiments that demonstrated high and stable efficiency rates, typically ranging from 90% to 95%, confirmed the rice protoplast transformation samples as qualified. Subsequently, these samples were advanced to the next-generation sequencing (NGS) experiments. As a result, a normalization strategy was not employed in the calculation of editing efficiency.

Rice stable transformation

As with our previous study (Lowder *et al.*, 2015), the cultivar *Japonica* Nipponbare was used for stable Agrobacterium-mediated transformation of rice. Briefly, rice seeds were dehulled, sterilized, and then cultured on solid N6-D medium (Zhong *et al.*, 2019). Precultured rice calli were transformed by inoculating Agrobacterium EHA105 carrying the recombinant expression vector. After the rice calli were co-cultured with Agrobacterium for 3 days in co-culture medium (Zhong *et al.*, 2019), the calli were washed with sterile water and transferred to N6-S medium for 2 weeks of selection (Zhong *et al.*, 2019). The newly grown calli were then transferred to RE-III medium and cultured for 2 weeks (Zhong *et al.*, 2019). Resistant calli were transferred to fresh RE-III medium every two weeks until regenerated plants were successfully obtained.

Detection and quantification of genome editing

The genome editing efficiency in rice protoplasts was assessed using Next-Generation Sequencing (NGS) of PCR amplicons. Forty-eight hours following the culture of transformed rice protoplasts, genomic DNA is extracted employing the CTAB method (Stewart Jr. and Via, 1993). Specific primers for amplifying the target gene were synthesized, each bearing a unique 6-base barcode sequence at the 5′-end to facilitate PCR and sample distinction (Zhong *et al.*, 2019). The success of amplification was verified through electrophoresis. Once amplification is confirmed for all samples, those with different barcodes are combined, gel-purified, and then dispatched to Novogene (Tianjin, China) for comprehensive sequencing via the Illumina HiSeqX platform. Sequencing data are analysed using CRISPR-Match (You *et al.*, 2018) and CrisprStitch (Han *et al.*, 2023). For plants derived from stable rice transformation, DNA was extracted from the T0 generation using the CTAB method (Stewart Jr. and Via, 1993), followed by target gene amplification and direct submission for Sanger sequencing. Sanger sequencing data were decoded using DSDecode (Xie *et al.*, 2017), facilitating the determination of the genotype for each T0 plant. Mutation types for individual plants were identified by comparing them with the reference genome, enabling the calculation of mutation and biallelic editing efficiencies.

Small RNA sequencing and mRNA transcriptome sequencing

The miRNA mutants and WT plants were chosen for mRNA transcriptome sequencing and small RNA sequencing. Plants were transferred to soil and grown in a growth chamber under long-day conditions (16-h light at 28 °C and 8-h dark at 22 °C) for 10 days. Then, leaf tissues were placed in self-sealing bags or 50-mL centrifuge tubes, rapidly frozen in liquid nitrogen, and the samples were sent to Biomarker Technologies Co. Ltd. (China) for RNA extraction, library construction, sequencing, and analysis. An Illumina HiSeq 2500 platform was used for mRNA transcriptome sequencing and small RNA sequencing. Data processing and analysis were conducted by Biomarker company using the BMKCloud service (http://www.biocloud.net/) (Zhou *et al.*, 2022).

RNA extraction and qRT-PCR

RNA extraction of WT and mutant seedings was carried out using the SteadyPure Plant RNA Isolation Kit (Accurate Biology, China) (Zhong *et al.*, 2023), and reverse transcription was performed with the HiScript III 1st Strand cDNA Synthesis Kit (Vazyme, China). Real-time qPCR was conducted using the ChamQ Universal SYBR qPCR Master Mix (Vazyme, China) following the manufacturer's instructions, with *OsActin* mRNA serving as an internal control. The relative levels of gene expression were calculated using the $2^{-\Delta\Delta Ct}$ method. Two biological replicates (two independent mutant leaves) were examined to ensure reproducibility (Zhong *et al.*, 2023). The experiment was performed three times independently, and similar results were obtained. All primers used in this study are listed in Table S6.

Whole-genome sequencing and data analysis

The rice plants identified through screening with Sanger sequencing were transplanted into growth chambers and cultivated at 28 °C with a 12-h light/12-h dark cycle. After 5–6 weeks of cultivation, 10 cm^2 leaf samples were collected from each plant, placed in self-sealing bags or 50 mL centrifuge tubes, rapidly frozen in liquid nitrogen, and stored in a −80 °C ultra-low temperature freezer. Two leaf samples were collected from each plant for backup. The DNA samples were sent to Molbreeding Company in Shijiazhuang, China, for library construction, using the HuaDa DNBSEQ-T7 platform for resequencing. The average sequencing data for each sample was 10 GB, with an average depth of approximately 20×−40×. We followed a similar WGS analysis pipeline as we previously demonstrated (Tang *et al.*, 2018). Briefly, the returned whole-genome

sequencing data underwent quality control and filtering using SKEWER software. The filtered data were aligned to the rice reference sequence using BWA software. Picard and Samtools software were employed to mark duplicate reads and generate BAM files. The GATK software was used for quality correction of insertions, deletions, and base substitutions. The analysis of whole-genome single nucleotide variations (SNVs) was conducted using LoFreq, MuTect2, and VarScan2 software. The analysis of whole-genome insertions and deletions (indels) was performed using MuTect2, VarScan2, and Pindel software. Bedtools and BCFtools were used to obtain files for SNVs and indels. CRISPR RGEN Tools were utilized to predict potential off-target sites in the rice genome. Data processing, analysis, and graphical representation were carried out using the R language and Python.

Data analysis

The data were analysed with the GraphPad Prism 9.0 software, and the figures were made using Adobe Photoshop and Adobe Illustrator software.

Acknowledgements

This research was supported by the STI 2030-Major Projects (2023ZD04074) to Y.Z., the National Natural Science Foundation of China (award no. 32270433, 32101205, and 32072045) to Y.Z., X.L.Z., and X.T. It is also supported by the NSF Plant Genome Research Program (IOS-2029889 and IOS-2132693) to Y.Q.

Conflict of interest

The authors declare no competing interests.

Author contributions

Y.Z. proposed the project and designed the experiments. Y.H. and Y.M. generated all the constructs. Y.H. did the rice protoplast transformation. Y.H., Yangshuo H., and T.Z. analysed the mutation frequencies in protoplasts. Y.H., Y.M., S.L. T.F., and Y.L. conducted rice stable transformation. Y.H. and Yangshuo H. analysed the mutation frequencies in rice stable T0 lines. Y.H. prepared rice seedling samples for WGS. Y.H. and Yangshuo H. performed WGS data analysis and generated the figures. Y.H. and Y.W. performed RNA-Seq and small RNA-Seq experiments. Yangshuo H. and T.Z. analysed the data. Y.Z., Y.Q., Y.H., and T.Z. analysed the data and wrote the paper with input from other authors. All authors read and approved the final version of the manuscript.

Data availability

The deep sequencing data sets generated from this study are available at the NCBI Sequence Read Archive under Bioproject PRJNA1050878, PRJNA1054222, and PRJNA1058782. pTrans_210d (#91109) has been previously deposited to Addgene. Backbone vectors are available from Addgene: pGEL846 (Addgene # 214236), pGEL847 (Addgene # 214237), pGEL848 (Addgene # 214238), and pGEL849 (Addgene # 214239).

References

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.

Cermak, T., Curtin, S.J., Gil-Humanes, J., Čegan, R., Kono, T.J.Y., Konečná, E., Belanto, J.J. *et al*. (2017) A multipurpose toolkit to enable advanced genome engineering in plants. *Plant Cell* **29**, 1196–1217.

Certo, M.T., Gwiazda, K.S., Kuhar, R., Sather, B., Curinga, G., Mandt, T., Brault, M. *et al*. (2012) Coupling endonucleases with DNA end-processing enzymes to drive gene disruption. *Nat. Methods* **9**, 973–975.

Chatterjee, P., Lee, J., Nip, L., Koseki, S.R.T., Tysinger, E., Sontheimer, E.J., Jacobson, J.M. *et al*. (2020) A Cas9 with PAM recognition for adenine dinucleotides. *Nat. Commun.* **11**, 2474.

Cheng, H.L., Lin, C.T., Huang, K.W., Wang, S., Lin, Y.T., Toh, S.I. and Hsiao, Y.Y. (2018) Structural insights into the duplex DNA processing of TREX2. *Nucleic Acids Res.* **46**, 12166–12176.

Cheng, Y., Zhang, Y., Li, G., Fang, H., Sretenovic, S., Fan, A., Li, J. *et al*. (2023) CRISPR-Cas12a base editors confer efficient multiplexed genome editing in rice. *Plant Commun.* **4**, 100601.

Chylinski, K., Makarova, K.S., Charpentier, E. and Koonin, E.V. (2014) Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Res.* **42**, 6091–6105.

Cui, Z., Tian, R., Huang, Z., Jin, Z., Li, L., Liu, J., Huang, Z. *et al*. (2022) FrCas9 is a CRISPR/Cas9 system with high editing efficiency and fidelity. *Nat. Commun.* **13**, 1425.

Curtin, S., Qi, Y., Peres, L.E.P., Fernie, A.R. and Zsogon, A. (2022) Pathways to de novo domestication of crop wild relatives. *Plant Physiol.* **188**, 1746–1756.

Fedorova, I., Arseniev, A., Selkova, P., Pobegalov, G., Goryanin, I., Vasileva, A., Musharova, O. *et al*. (2020) DNA targeting by *Clostridium cellulolyticum* CRISPR-Cas9 Type II-C system. *Nucleic Acids Res.* **48**, 2026–2034.

Gaillochet, C., Pena Fernandez, A., Goossens, V., D'Halluin, K., Drozdzecki, A., Shafie, M., Van Duyse, J. *et al*. (2023) Systematic optimization of Cas12a base editors in wheat and maize using the ITER platform. *Genome Biol.* **24**, 6.

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

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

Han, Y., Liu, G., Wu, Y., Bao, Y., Zhang, Y. and Zhang, T. (2023) CrisprStitch: Fast evaluation of the efficiency of CRISPR editing systems. *Plant Commun.* **5**, 100783.

Harrington, L.B., Ma, E., Chen, J.S., Witte, I.P., Gertz, D., Paez-Espino, D., al-Shayeb, B. *et al*. (2020) A scoutRNA is required for some type V CRISPR-Cas systems. *Mol. Cell* **79**, 416–424.e5.

Hirano, S., Abudayyeh, O.O., Gootenberg, J.S., Horii, T., Ishitani, R., Hatada, I., Zhang, F. *et al*. (2019) Structural basis for the promiscuous PAM recognition by *Corynebacterium diphtheriae* Cas9. *Nat. Commun.* **10**, 1968.

Hou, Z., Zhang, Y., Propson, N.E., Howden, S.E., Chu, L.F., Sontheimer, E.J. and Thomson, J.A. (2013) Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc. Natl. Acad. Sci. USA* **110**, 15644–15649.

Hu, X., Wang, C., Fu, Y., Liu, Q., Jiao, X. and Wang, K. (2016) Expanding the range of CRISPR/Cas9 genome editing in rice. *Mol. Plant* **9**, 943–945.

Hu, Z., Wang, S., Zhang, C., Gao, N., Li, M., Wang, D., Wang, D. *et al*. (2020) A compact Cas9 ortholog from *Staphylococcus auricularis* (SauriCas9) expands the DNA targeting scope. *PLoS Biol.* **18**, e3000686.

Jiang, W., Bikard, D., Cox, D., Zhang, F. and Marraffini, L.A. (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **31**, 233–239.

Jiao, Y., Wang, Y., Xue, D., Wang, J., Yan, M., Liu, G., Dong, G. *et al*. (2010) Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. *Nat. Genet.* **42**, 541–544.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821.

Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Topkar, V.V., Nguyen, N.T., Zheng, Z., Gonzales, A.P.W. *et al*. (2015) Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485.

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

Kong, X., Zhang, H., Li, G., Wang, Z., Kong, X., Wang, L., Xue, M. *et al*. (2023) Engineered CRISPR-OsCas12f1 and RhCas12f1 with robust activities and expanded target range for genome editing. *Nat. Commun.* **14**, 2046.

Li, G., Sretenovic, S., Eisenstein, E., Coleman, G. and Qi, Y. (2021) Highly efficient C-to-T and A-to-G base editing in a Populus hybrid. *Plant Biotechnol. J.* **19**, 1086–1088.

Li, J.F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M. *et al*. (2013) Multiplex and homologous recombination-mediated genome editing in Arabidopsis and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.* **31**, 688–691.

Li, S., Liu, L., Sun, W., Zhou, X. and Zhou, H. (2022) A large-scale genome and transcriptome sequencing analysis reveals the mutation landscapes induced by high-activity adenine base editors in plants. *Genome Biol.* **23**, 51.

Liu, G., Yin, K., Zhang, Q., Gao, C. and Qiu, J.L. (2019) Modulating chromatin accessibility by transactivation and targeting proximal dsgRNAs enhances Cas9 editing efficiency in vivo. *Genome Biol.* **20**, 145.

Liu, S., Sretenovic, S., Fan, T., Cheng, Y., Li, G., Qi, A., Tang, X. *et al*. (2022) Hypercompact CRISPR-Cas12j2 (CasPhi) enables genome editing, gene activation, and epigenome editing in plants. *Plant Commun.* **3**, 100453.

Lowder, L.G., Zhang, D., Baltes, N.J., Paul, J.W., III, Tang, X., Zheng, X., Voytas, D.F. *et al*. (2015) A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. *Plant Physiol.* **169**, 971–985.

Lv, P., Su, F., Chen, F., Yan, C., Xia, D., Sun, H., Li, S. *et al*. (2023) Genome editing in rice using CRISPR/Cas12i3. *Plant Biotechnol. J.* **22**, 379–385.

Mazur, D.J. and Perrino, F.W. (1999) Identification and expression of the TREX1 and TREX2 cDNA sequences encoding mammalian 3′→5′ exonucleases. *J. Biol. Chem.* **274**, 19655–19660.

Ming, M., Ren, Q., Pan, C., He, Y., Zhang, Y., Liu, S., Zhong, Z. *et al*. (2020) CRISPR-Cas12b enables efficient plant genome engineering. *Nat. Plants* **6**, 202–208.

Molla, K.A., Sretenovic, S., Bansal, K.C. and Qi, Y. (2021) Precise plant genome editing using base editors and prime editors. *Nat. Plants* **7**, 1166–1187.

Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., Mochizuki, M. *et al*. (2016) Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **353**, aaf8729.

Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T. *et al*. (2018) Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* **361**, 1259–1262.

Qin, R., Li, J., Li, H., Zhang, Y., Liu, X., Miao, Y., Zhang, X. *et al*. (2019) Developing a highly efficient and wildly adaptive CRISPR-SaCas9 toolset for plant genome editing. *Plant Biotechnol. J.* **17**, 706–708.

Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B. *et al*. (2015) In vivo genome editing using Staphylococcus aureus Cas9. *Nature* **520**, 186–191.

Randall, L.B., Sretenovic, S., Wu, Y., Yin, D., Zhang, T., Eck, J.V. and Qi, Y. (2021) Genome- and transcriptome-wide off-target analyses of an improved cytosine base editor. *Plant Physiol.* **187**, 73–87.

Ren, B., Liu, L., Li, S., Kuang, Y., Wang, J., Zhang, D., Zhou, X. *et al*. (2019) Cas9-NG greatly expands the targeting scope of the genome-editing toolkit by recognizing NG and other atypical PAMs in rice. *Mol. Plant* **12**, 1015–1026.

Ren, J., Meng, X., Hu, F., Liu, Q., Cao, Y., Li, H., Yan, C. *et al*. (2021a) Expanding the scope of genome editing with SpG and SpRY variants in rice. *Sci. China Life Sci.* **64**, 1784–1787.

Ren, Q., Sretenovic, S., Liu, G., Zhong, Z., Wang, J., Huang, L., Tang, X. *et al*. (2021b) Improved plant cytosine base editors with high editing activity, purity, and specificity. *Plant Biotechnol. J.* **19**, 2052–2068.

Ren, Q., Sretenovic, S., Liu, S., Tang, X., Huang, L., He, Y., Liu, L. *et al*. (2021c) PAM-less plant genome editing using a CRISPR-SpRY toolbox. *Nat. Plants* **7**, 25–33.

Richter, M.F., Zhao, K.T., Eton, E., Lapinaite, A., Newby, G.A., Thuronyi, B.W., Wilson, C. *et al*. (2020) Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.* **38**, 883–891.

Rodriguez-Leal, D., Lemmon, Z.H., Man, J., Bartlett, M.E. and Lippman, Z.B. (2017) Engineering quantitative trait variation for crop improvement by genome editing. *Cell* **171**, 470–480.e8.

Sander, J.D. and Joung, J.K. (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* **32**, 347–355.

Shao, Y., Zhou, H.Z., Wu, Y., Zhang, H., Lin, J., Jiang, X., He, Q. *et al*. (2019) OsSPL3, an SBP-domain protein, regulates crown root development in rice. *Plant Cell* **31**, 1257–1275.

Sretenovic, S., Green, Y., Wu, Y., Cheng, Y., Zhang, T., van Eck, J. and Qi, Y. (2023a) Genome- and transcriptome-wide off-target analyses of a high-efficiency adenine base editor in tomato. *Plant Physiol.* **193**, 291–303.

Sretenovic, S., Tang, X., Ren, Q., Zhang, Y. and Qi, Y. (2023b) PAM-less CRISPR-SpRY genome editing in plants. *Methods Mol. Biol.* **2653**, 3–19.

Sretenovic, S., Yin, D., Levav, A., Selengut, J.D., Mount, S.M. and Qi, Y. (2021) Expanding plant genome-editing scope by an engineered iSpyMacCas9 system that targets A-rich PAM sequences. *Plant Commun.* **2**, 100101.

Stewart, C.N., Jr. and Via, L.E. (1993) A rapid CTAB DNA isolation technique useful for RAPD fingerprinting and other PCR applications. *BioTechniques* **14**, 748–750.

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, X., Lowder, L.G., Zhang, T., Malzahn, A.A., Zheng, X., Voytas, D.F., Zhong, Z. *et al*. (2017) A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nat. Plants* **3**, 17103.

Tang, X. and Zhang, Y. (2023) Beyond knockouts: fine-tuning regulation of gene expression in plants with CRISPR-Cas-based promoter editing. *New Phytol.* **239**, 868–874.

Tseng, B.S., Huang, C.C., King, Y.C., Wu, M.T., Hsieh, C.H., Hsieh, K.T., Hsing, Y.I. *et al*. (2023) Hydrogen peroxide regulates the Osa-miR156-OsSPL2/OsTIFY11b module in rice. *Plant Cell Environ.* **46**, 2507–2522.

Walton, R.T., Christie, K.A., Whittaker, M.N. and Kleinstiver, B.P. (2020) Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* **368**, 290–296.

Weiss, T., Crisp, P.A., Rai, K.M., Song, M., Springer, N.M. and Zhang, F. (2022) Epigenetic features drastically impact CRISPR-Cas9 efficacy in plants. *Plant Physiol.* **190**, 1153–1164.

Wu, W.Y., Mohanraju, P., Liao, C., Adiego-Pérez, B., Creutzburg, S.C.A., Makarova, K.S., Keessen, K. *et al*. (2022a) The miniature CRISPR-Cas12m effector binds DNA to block transcription. *Mol. Cell* **82**, e7.

Wu, Y., Ren, Q., Zhong, Z., Liu, G., Han, Y., Bao, Y., Liu, L. *et al*. (2022b) 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., 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.

Xie, X., Ma, X., Zhu, Q., Zeng, D., Li, G. and Liu, Y.G. (2017) CRISPR-GE: a convenient software toolkit for CRISPR-based genome editing. *Mol. Plant* **10**, 1246–1249.

Xu, R., Qin, R., Xie, H., Li, J., Liu, X., Zhu, M., Sun, Y. *et al*. (2022) Genome editing with type II-C CRISPR-Cas9 systems from *Neisseria meningitidis* in rice. *Plant Biotechnol. J.* **20**, 350–359.

Xu, Y., Meng, X., Wang, J., Qin, B., Wang, K., Li, J., Wang, C. *et al*. (2020) ScCas9 recognizes NNG protospacer adjacent motif in genome editing of rice. *Sci. China Life Sci.* **63**, 450–452.

Xu, Z., Kuang, Y., Ren, B., Yan, D., Yan, F., Spetz, C., Sun, W. *et al*. (2021) SpRY greatly expands the genome editing scope in rice with highly flexible PAM recognition. *Genome Biol.* **22**, 6.

You, Q., Zhong, Z., Ren, Q., Hassan, F., Zhang, Y. and Zhang, T. (2018) CRISPRMatch: an automatic calculation and visualization tool for high-throughput CRISPR genome-editing data analysis. *Int. J. Biol. Sci.* **14**, 858–862.

Yu, H., Lin, T., Meng, X., du, H., Zhang, J., Liu, G., Chen, M. *et al*. (2021) A route to de novo domestication of wild allotetraploid rice. *Cell* **184**, 1156–1170.e14.

Zetsche, B., Abudayyeh, O.O., Gootenberg, J.S., Scott, D.A. and Zhang, F. (2020) A survey of genome editing activity for 16 Cas12a orthologs. *Keio J. Med.* **69**, 59–65.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E. *et al*. (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771.

Zhang, H., Kong, X., Xue, M., Hu, J., Wang, Z., Wei, Y., Wang, H. *et al*. (2023) An engineered xCas12i with high activity, high specificity, and broad PAM range. *Protein Cell* **14**, 538–543.

Zhang, Y., Ren, Q., Tang, X., Liu, S., Malzahn, A.A., Zhou, J., Wang, J. *et al*. (2021) Expanding the scope of plant genome engineering with Cas12a orthologs and highly multiplexable editing systems. *Nat. Commun.* **12**, 1944.

Zhong, Z., Liu, G., Tang, Z., Xiang, S., Yang, L., Huang, L., He, Y. *et al*. (2023) Efficient plant genome engineering using a probiotic sourced CRISPR-Cas9 system. *Nat. Commun.* **14**, 6102.

Zhong, Z., Sretenovic, S., Ren, Q., Yang, L., Bao, Y., Qi, C., Yuan, M. *et al*. (2019) Improving plant genome editing with high-fidelity xCas9 and non-canonical PAM-targeting Cas9-NG. *Mol. Plant* **12**, 1027–1036.

Zhong, Z., Zhang, Y., You, Q., Tang, X., Ren, Q., Liu, S., Yang, L. *et al*. (2018) Plant genome editing using FnCpf1 and LbCpf1 nucleases at redefined and altered PAM sites. *Mol. Plant* **11**, 999–1002.

Zhou, C., Sun, Y., Yan, R., Liu, Y., Zuo, E., Gu, C., Han, L. *et al*. (2019) Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature* **571**, 275–278.

Zhou, J., Liu, G., Zhao, Y., Zhang, R., Tang, X., Li, L., Jia, X. *et al*. (2023) An efficient CRISPR-Cas12a promoter editing system for crop improvement. *Nat Plants* **9**, 588–604.

Zhou, J., Zhang, R., Jia, X., Tang, X., Guo, Y., Yang, H., Zheng, X. *et al*. (2022) CRISPR-Cas9 mediated OsMIR168a knockout reveals its pleiotropy in rice. *Plant Biotechnol. J.* **20**, 310–322.

Zhu, Y., Gao, A., Zhan, Q., Wang, Y., Feng, H., Liu, S., Gao, G. *et al*. (2019) Diverse mechanisms of CRISPR-Cas9 inhibition by type IIC anti-CRISPR proteins. *Mol. Cell* **74**, 296–309.e7.

Supporting information

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

Figure S1 The gene structures of the FrCas9 and SpCas9 CRISPR systems.
Figure S2 Genotype of FrCas9-mediated singular genome editing at the *OsEPFP9*-sgR01 and *OsGn1a*-sgR01 sites in rice T0 lines.
Figure S3 Genotype of FrCas9-mediated singular genome editing at the *OsGS3*-sgR01 and *OsROC5*-sgR01 sites in rice T0 lines.
Figure S4 Genotype of FrCas9-mediated singular genome editing at the *OsGBSSI*-sgR01 and *OsDEP1*-sgR01 sites in rice T0 lines.
Figure S5 TREX2-FrCas9 induces larger deletions in 18 rice loci.
Figure S6 Genotype of TREX2-FrCas9-mediated genome editing at *OsMIR156j* in rice T0 lines.
Figure S7 Expression analysis of pri-MIR156j in the *osmir156j*-m10 and *osmir156j*-m12 mutants.
Figure S8 The expression profiles of the *OsSPL* gene, respectively, based on the mRNA transcriptome data.
Figure S9 Indel byproduct frequency by FrCas9 CBE in rice protoplasts.
Figure S10 Indel byproduct frequency by FrCas9 ABE in rice protoplasts.
Figure S11 Off-target analysis of FrCas9 editors.
Table S1 Information on RNA-Seq.
Table S2 Information on small RNA-Seq.
Table S3 Genotype of edited rice T0 lines used for whole-genome sequencing.
Table S4 Information on whole-genome sequencing.
Table S5 Target sites used in this study.
Table S6 Oligos and sequences used in this study.