nature plants

BRIEF COMMUNICATION

https://doi.org/10.1038/s41477-020-0614-6

CRISPR–Cas12b enables efficient plant genome engineering

Meiling Ming[1,2,5], Qiurong Ren[3,5], Changtian Pan[1,5], Yao He[3,5], Yingxiao Zhang[1,5], Shishi Liu[3], Zhaohui Zhong[3], Jiaheng Wang[3], Aimee A. Malzahn[1], Jun Wu[2], Xuelian Zheng[3], Yong Zhang[3]✉ and Yiping Qi[1,4]✉

Clustered regularly interspaced short palindromic repeats (CRISPR)-Cas12b is a newly emerged genome engineering system. Here, we compared Cas12b from *Alicyclobacillus acidoterrestris* (Aac), *Alicyclobacillus acidiphilus* (Aa), *Bacillus thermoamylovorans* (Bth) and *Bacillus hisashii* (Bh) for genome engineering in rice, an important crop. We found AaCas12b was more efficient than AacCas12b and BthCas12b for targeted mutagenesis, which was further demonstrated in multiplexed genome editing. We also engineered the Cas12b systems for targeted transcriptional repression and activation. Our work establishes Cas12b as the third promising CRISPR system, after Cas9 and Cas12a, for plant genome engineering.

In recent years, CRISPR–Cas9 (ref. [1]) and Cas12a[2], as RNA-guided endonuclease systems, have become leading sequence-specific nucleases (SSNs) in plant genome engineering[3]. Cas12b (formerly C2c1), a class 2 type V-B CRISPR system[4], was recently demonstrated as a new SSN for mammalian genome editing[5,6]. Similar to Cas12a (formerly Cpf1, a class 2 type V-A system), Cas12b prefers T-rich protospacer adjacent motifs (PAMs) and generates staggered ends of DNA double-strand breaks[5–9]. Similar to Cas9 (a class 2 type II system), Cas12b requires a CRISPR RNA (crRNA) and a trans-activating crRNA, which can be combined as a single guide RNA (sgRNA), for DNA targeting[7–9]. By contrast, Cas12a only requires a crRNA. Hence, Cas12b is more amendable than Cas12a with versatile guide RNA engineering[3]. In addition, Cas12b is smaller than Cas9 and Cas12a in protein size[5,6]. In human and mouse cells, AaCas12b can barely tolerate single base pair mismatches in the protospacer, suggesting it has high targeting specificity[5]. For these reasons, it is desirable to develop Cas12b systems for plant genome engineering.

Structures for DNA targeting complexes of AacCas12b and BthCas12b have been recently resolved[7–9]. We decided to test AacCas12b, AaCas12b and BthCas12b for their capability in plant genome editing. Since AaCas12b shares high sequence identity with AacCas12b[7] (Supplementary Fig. 1), the AacCas12b sgRNA scaffold was used for both AacCas12b and AaCas12b[7,8]. Similarly, a BthCas12b sgRNA scaffold was used for BthCas12b[9]. These Cas12b DNA coding sequences were codon-optimized for rice, a major crop and test platform in this study. We adopted the dual Polymerase II (Pol II) promoter expression system and hammerhead virus–hepatitis delta virus dual ribozyme guide RNA processing system that we established for CRISPR–Cas12a[10,11] (Fig. 1a). Previous in vitro assays established PAMs as TTN (N = A, T, G, C) for AacCas12b[5] and ATTN for BthCas12b[12]. We targeted two sites in *OsEPFL9* and *OsGS3* with GTTG and ATTC PAMs, respectively. To quantify the editing efficiencies of Cas12b nucleases, expression vectors were transfected into rice protoplasts. AaCas12b resulted in an editing efficiency over 10% at both sites, higher than AacCas12b (~5%) (Fig. 1b). BthCas12b displayed very low editing efficiency (Fig. 1b). AaCas12b, AacCas12b and BthCas12b mainly generated 4–14 base pair (bp) deletions (Fig. 1c,d and Supplementary Fig. 2), which are larger than those induced by Cas9 (1–3 bp)[13]. These deletions occurred about 12–24 nucleotides distal to the PAM sites (Fig. 1e,f and Supplementary Figs. 2 and 3), consistent with the staggered double-strand breaks generated in this region[5,6]. Targeting an additional site in *OsPDS* with AacCas12b further confirmed this editing pattern (Supplementary Fig. 4).

To further investigate the PAM requirements for AacCas12b and AaCas12b *in planta*, we targeted a series of VTTV (V = A, C, G) PAM sites and assessed editing activity in rice protoplasts. While both AacCas12b and AaCas12b showed editing activity at five out of six ATTV sites, AaCas12b is generally more efficient and resulted in over 50% mutation frequencies at ATTA-01 and ATTC-01 sites (Fig. 1g and Supplementary Fig. 5). Among two additional GTTG PAM sites, both AacCas12b and AaCas12b resulted in high editing efficiency (50%–60%) at one site (GTTG-01) but failed at the other site (GTTG-02) (Fig. 1g and Supplementary Fig. 5). Further testing suggested AaCas12b could edit CTTG and GTTC PAM sites (Fig. 1h and Supplementary Fig. 6). However, both Cas12b variants largely failed at an additional three CTTG and two GTTC PAM sites, as well as three CTTC and two GTTA PAM sites (Supplementary Fig. 7). Unlike Cas12a[10], AacCas12b and AaCas12b could barely edit six VTTTV PAM sites tested (Supplementary Fig. 8). Interestingly, AaCas12b could edit a TTTTV PAM site with ~20% mutation frequency (Fig. 1h and Supplementary Fig. 6). Together, our data demonstrate that AaCas12b and AacCas12b are potent SSNs for targeted mutagenesis in rice and they generally recognize VTTV PAMs, with more preference for ATTV and GTTG PAMs. Our observation is largely consistent with the observations of PAM requirements for Cas12b orthologs in human cells[5,6].

Initial comparison of three Cas12b orthologs suggested that AaCas12b is superior to AacCas12b and BthCas12b for targeted mutagenesis in rice. We assessed targeting specificity of AaCas12b

[1]Department of Plant Science and Landscape Architecture, University of Maryland, College Park, MD, USA. [2]Centre of Pear Engineering Technology Research, State Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing, China. [3]Department of Biotechnology, School of Life Science and Technology, Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu, China. [4]Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, MD, USA. [5]These authors contributed equally: Meiling Ming, Qiurong Ren, Changtian Pan, Yao He, Yingxiao Zhang. ✉e-mail: zhangyong916@uestc.edu.cn; yiping@umd.edu

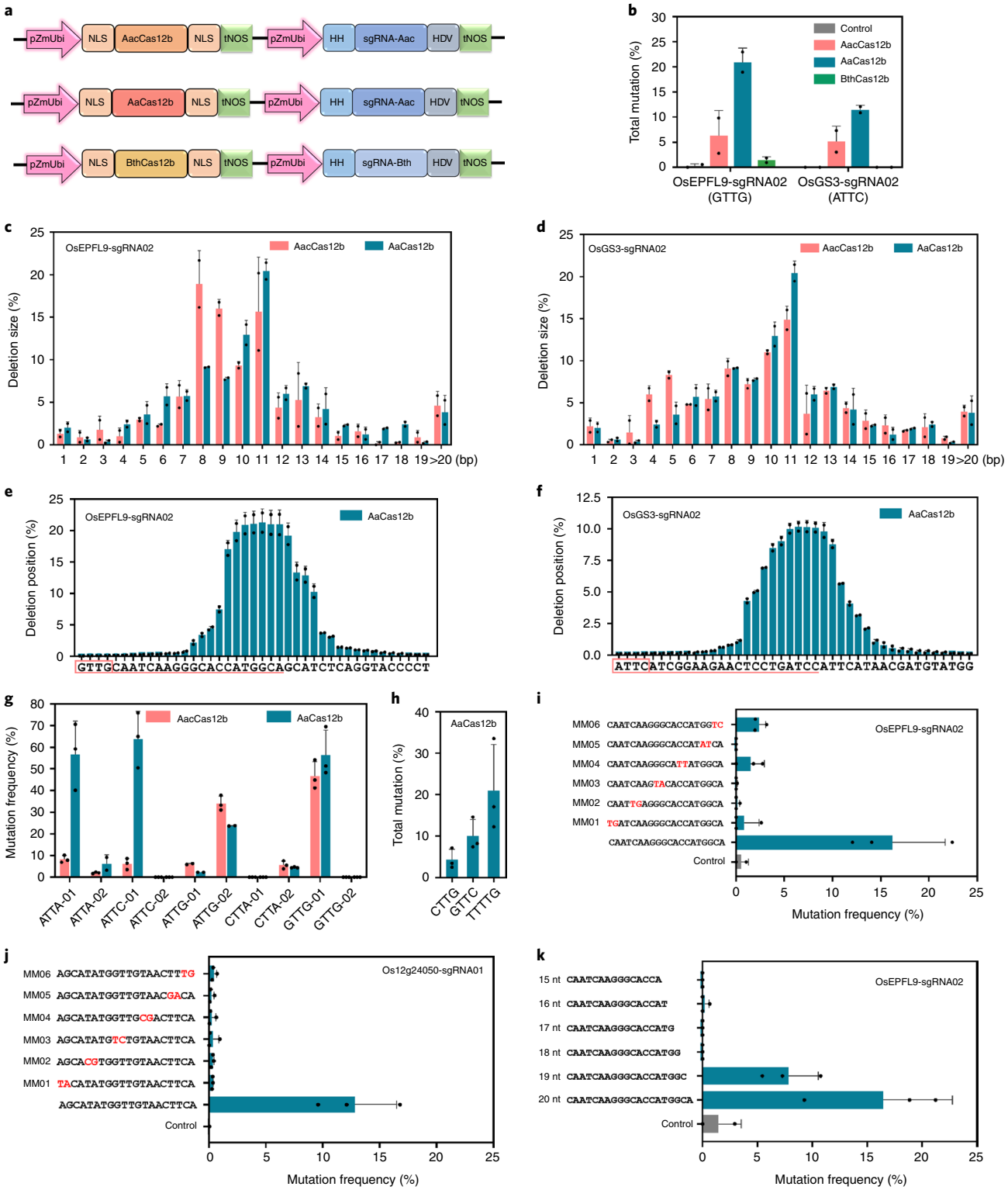


Fig. 1 | Comprehensive analysis of three CRISPR–Cas12b systems for genome editing in rice protoplasts. a, Illustration of the dual Pol II promoter system for expression of Cas12b and sgRNA. Note the sgRNA is flanked by hammerhead (HH) and hepatitis delta virus (HDV) ribozymes for precise processing. NLS, nuclear localization signal; tNOS, NOS terminator. **b**, Comparison of mutation frequencies by AacCas12b, AaCas12b and BthCas12b at two target sites. **c**,**d**, Comparison of deletion sizes by AacCas12b and AaCas12b at the OsEPFL9-sgRNA02 site (**c**) and the OsGS3-sgRNA02 site (**d**). **e**,**f**, Comparison of deletion position by AaCas12b at the OsEPFL9-sgRNA02 site (**e**) and OsGS3-sgRNA02 site (**f**). PAM and protospacer sequences are circled and underlined, respectively. **g**, Comparison of mutation frequencies by AacCas12b and AaCas12b at 10 sites with ATTV, CTTA and GTTG PAMs. **h**, Targeted mutagenesis by AaCas12b at an additional three PAM sites. **i**,**j**, Off-targeting analysis with mismatch (MM) sgRNAs at the OsEPFL9-sgRNA02 site (**i**) and Os12g24050-sgRNA01 site (**j**) by AaCas12b. **k**, Comparison of protospacer length for targeted mutagenesis at the OsEPFL9-sgRNA02 site by AaCas12b. Data of **b**–**h** were generated from high-throughput sequencing while data of **i**,**j** were generated from RFLP analysis. Mean values of two biological replicates were shown for **b**–**f**. Mean values of three biological replicates were shown for **g**–**k**. Error bars represent standard deviations.

by using six crRNA protospacer sequences of OsEPFL9-sgRNA02 that carry double mismatch nucleotides (at positions 1–2, 5–6, 9–10, 13–14, 17–18 and 19–20). These six constructs were compared with the on-target control construct in rice protoplasts. The mutation frequency data suggested that all these mismatched nucleotides had completely abolished editing activity at the target site (Fig. 1i and Supplementary Fig. 9). Similar results were obtained by targeting an independent site with Os12g24050-sgRNA01 (Fig. 1j), suggesting that AaCas12b is a highly specific SSN in rice cells. Interestingly, a recent study in human and mouse cells suggests that AaCas12b, unlike Cas9 and Cas12a, could barely tolerate single base mismatches at nearly every position of the protospacer, supporting its high specificity[5]. We further shortened the length of the protospacer of OsEPFL9-sgRNA02 and found that AaCas12b completely lost editing activity with protospacers of 18 nucleotides and shorter (Fig. 1k and Supplementary Fig. 9). While further study is warranted, this result is in sharp contrast to Cas9 and Cas12a, which generally still possess nuclease activity with 17–18-nucleotide protospacers[2,14]. Together, our data suggest that AaCas12b is a highly specific SSN for plant genome editing.

We next sought to generate rice mutants by Cas12b. Both AacCas12b and AaCas12b constructs targeting the OsEPFL9-sgRNA02 site were transformed into rice calli by *Agrobacterium*. Analysis of 22 individual T0 transgenic lines for AacCas12b revealed eight lines which carried monoallelic mutations at the target site, representing a 36.4% mutation rate (Fig. 2a). Consistent with the protoplast data, AaCas12b had a higher mutation rate of 54.2% as 13 out of 24 T0 lines were mutants and 6 lines carried biallelic mutations (Fig. 2b). The mutations in these edited lines were predominantly large deletions (Fig. 2a, b and Supplementary Fig. 10). These results demonstrated that both AacCas12b and AaCas12b can effectively generate stable mutants in rice.

A major advantage of the CRISPR system is its flexibility of multiplexing. We constructed a multiplexed Cas12b system based on dual Pol II promoters and a hammerhead–sgRNA–hepatitis delta virus array[10] (Fig. 2c). We decided to simultaneously target three rice genes with three sgRNAs: OsROC5-sgRNA02, OsEPFL9-sgRNA02 and OsGS3-sgRNA02. Two multiplexing constructs based on AacCas12b and AaCas12b were made for rice stable transformation. For each construct, we analysed 24 independent T0 lines. For AacCas12b, 1 line (Line 17) carried a monoallelic mutation at the OsROC5-sgRNA02 site; 12 (50%) lines had mutations at the OsEPFL9-sgRNA02 site and 3 lines had biallelic mutations; 6 (25%) lines had mutations at the OsGS3-sgRNA02 site and none carried biallelic mutations (Fig. 2d). Among them, 4 (lines 1, 17, 20 and 22) are double mutants (Fig. 2d). These mutations were first identified by restriction fragment length polymorphism (RFLP) assays and later confirmed by Sanger sequencing (Supplementary Fig. 11). For AaCas12b, none of the 24 T0 plants assayed carried mutations at the OsROC5-sgRNA02 site, consistent with the low editing activity for this sgRNA in protoplasts (Fig. 1h). However, AaCas12b resulted in very high mutation rates at both *OsEPFL9* and *OsGS3* genes: at the OsEPFL9-sgRNA02 site, 16 (66.7%) T0 lines were mutants and 7 lines had biallelic mutations; at the OsGS3-sgRNA02 site, 17 (70.85%) T0 lines were mutants and 11 lines had biallelic mutations (Fig. 2e). Impressively, 16 lines were double mutants and 7 were biallelic double mutants (Fig. 2e). These mutations, including a 118-bp large deletion, have been further validated by Sanger sequencing (Supplementary Fig. 12). To assess off-target effects in T0 lines, we randomly selected two double mutants each generated by AacCas12b and AaCas12b. In both cases, sequencing of 7–8 top putative off-target sites of OsEPFL9-sgRNA02 and OsGS3-sgRNA02 revealed no off-target mutations (Supplementary Figs. 13 and 14). Taken together, we have successfully demonstrated multiplexed genome editing by generating combinational mutants with highly specific AacCas12b and AaCas12b.

We previously established CRISPR interference (CRISPRi) systems in plants based on Cas9 and Cas12a, which recognize NGG (for SpCas9) and TTTV (for AsCas12a and LbCas12a) PAMs, respectively[10,15]. As Cas12b orthologs have different PAM requirements, repurposing them for CRISPRi may expand the targeting range for plant transcriptional repression. We introduced single amino acid mutations at RuvC-I (D570A), RuvC-II (E848A) and RuvC-III (D977A) in AacCas12b and the corresponding mutations in AaCas12b and BthCas12b (Fig. 3a). Assessment of these protein variants of AacCas12b and AaCas12b in rice protoplasts revealed that they indeed lost nuclease activity (Fig. 3b,c and Supplementary Fig. 15). We chose three of these deactivated Cas12b (dCas12b) proteins, AacCas12b-D570A, AaCas12b-D570A and BthCas12b-D573A, to test CRISPRi in rice cells. We targeted the rice gene Os04g39780 by focusing on three PAMs: ATTC, CTTG and GTTG. For each PAM, we designed three sgRNAs that target either the promoter or the coding sequence (Fig. 3d and Supplementary Fig. 16a). The resulting 27 CRISPRi constructs were tested in rice protoplasts and the target gene expression was quantified by quantitative PCR with reverse transcription (qRT–PCR). Three out of nine dBthCas12b constructs resulted in transcriptional repression (Fig. 3e–g), indicating BthCas12b could not robustly bind to the target DNA, consistent with the genome editing data (Fig. 1b). Both dAacCas12b and dAaCas12b induced transcriptional repression at nearly every target site with variable repression levels (25–75%) (Fig. 3e–g). Interestingly, targeted binding of dCas12b to the promoter region and the coding sequence can both robustly repress the target gene expression (Fig. 3e–g). We further fused three copies of SRDX repressor domain to the C termini of the dCas12b proteins and generated three synthetic transcriptional repressors (Supplementary Fig. 16b). By targeting the CTTG PAM sites with the same sgRNAs, we found these dCas12b-SRDX repressors resulted in comparable levels of gene repression (Supplementary Fig. 16c) to dCas12b (Fig. 3f). The data suggest that the CRISPRi effects are predominantly contributed to by transcription interference through dCas12b binding, rather than through chromatin modifications by the SRDX repressor.

To our knowledge, to date there has been no successful report of Cas12a transcriptional activation systems in plants. We previously reported an improved Cas9-based transcriptional activation system that used an engineered sgRNA2.0 scaffold with MS2 aptamers for recruiting transcriptional activators[16]. Such guide RNA engineering could be applied to Cas12b, but not Cas12a because Cas12a uses very short crRNAs that are incompatible with MS2 aptamer insertions. To establish efficient Cas12b-based transcriptional activation systems, we first sought to engineer the sgRNA scaffold to improve the overall editing efficiency. We tested AaCas12b genome editing with the artsgRNA13 scaffold and three engineered artsgRNA13 scaffolds with 1–2 MS2 insertions (Supplementary Fig. 17a)[17]. No editing activity was detected with these new scaffolds in rice protoplasts (Supplementary Fig. 17b). However, AaCas12b, when coupled with the scaffolds Aa1.2 and Aa3.8 (ref. [5]), showed comparable editing efficiencies with the Aac scaffold at four independent target sites (Supplementary Fig. 18a,c,d). Recently, an engineered Cas12b from *Bacillus hisashii* (Bh), BhCas12b-v4, was reported for genome editing in human cells[6]. We compared a rice codon-optimized BhCas12b-v4 with our AaCas12b systems and found AaCas12b showed equivalent or even better editing efficiency than BhCas12b-v4 (Supplementary Fig. 18a–d). We continued our focus on AaCas12b and sought to use engineered sgRNAs to recruit more activators for developing Cas12b-based transcriptional activation systems. Four sgRNA scaffolds (Aac.3, Aa1.2.3, Aa3.8.3 and Aa3.8.4) that contained one MS2 aptamer near the 3' end were first tested for genome editing (Supplementary Fig. 19a). Although all four modified sgRNA scaffolds led to detectable editing activities at two target sites in rice protoplasts, Aa3.8.4 had the highest editing

a

AacCas12b reagents	Tested T0 lines	Mutated T0 lines (number; percentage of total)	Biallelic mutation lines (number; percentage of total)
OsEPFL9-sgRNA02-L10 (pLR1380)	22	8; 36.4%	0; 0%

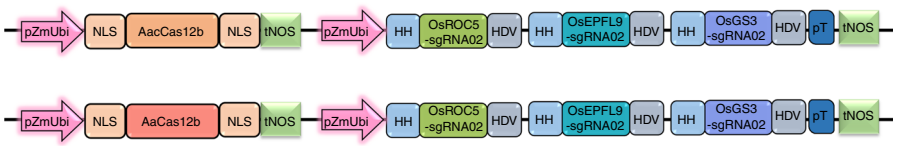
b

AaCas12b reagents	Tested T0 lines	Mutated T0 lines (number; percentage of total)	Biallelic mutation lines (number; percentage of total)
OsEPFL9-sgRNA02-L10 (pLR1687)	24	13; 54.2%	6; 25.0%

```
pLR1380: AacCas12b+OsEPFL9-sgRNA02-L10
          GCTCGTTGCAATCAAGGGCACCATGGCAGCATCTCAGGTACCCCT WT(OsEPFL9)
pLR1380-01
Allele 1: GCTCGTTGCAATCAAGGGCACCATGGCAGCATCTCAGGTACCCCT WT
Allele 2: GCTCGTTGCAATCAAGGGCACCAT--------CTCAGGTACCCCT -8bp
pLR1380-03
Allele 1: GCTCGTTGCAATCAAGGGCACCATGGCAGCATCTCAGGTACCCCT WT
Allele 2: GCTCGTTGCAATCAAGGGCACCAT----------CAGGTACCCCT -10bp
pLR1380-09
Allele 1: GCTCGTTGCAATCAAGGGCACCATGGCAGCATCTCAGGTACCCCT WT
Allele 2: GCTCGTTGCAATC--------------------CTCAGGTACCCCT -19bp
pLR1380-12
Allele 1: GCTCGTTGCAATCAAGGGCACCATGGCAGCATCTCAGGTACCCCT WT
Allele 2: GCTCGTTGCAATCAAGGGCACCA--GCAGCATCTCAGGTACCCCT -2bp
pLR1380-13
Allele 1: GCTCGTTGCAATCAAGGGCACCATGGCAGCATCTCAGGTACCCCT WT
Allele 2: GCTCGTTGCAATCAAGGGCACCA--GCAGCATCTCAGGTACCCCT -2bp
```

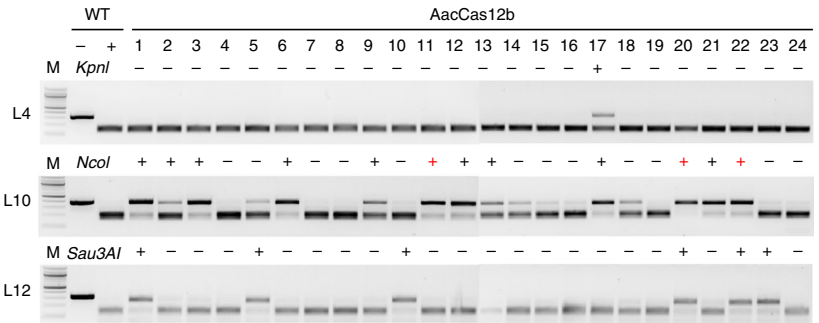
```
pLR1687: AaCas12b+OsEPFL9-sgRNA02-L10
          GCTCGTTGCAATCAAGGGCACCATGGCAGCATCTCAGGTACCCCT WT(OsEPFL9)
pLR1687-03
Allele 1: GCTCGTTGCAATCAAGGGCAC------------TCAGGTACCCCT -12bp
Allele 2: GCTCGTTGCAATCAAGGGCACCAT-----------GGTACCCCT -12bp
pLR1687-06
Allele 1: GCTCGTTGCAATC--------------------TCAGGTACCCCT -20bp
Allele 2: GCTCGTTGCAATCAAGGGCAC--------CATCTCAGGTACCCCT -8bp
pLR1687-09
Allele 1: GCTCGTTGCAATCAAGGGC-------------CTCAGGTACCCCT -13bp
Allele 2: GCTCGTTGCAATCAAGGGCAC--------CATCTCAGGTACCCCT -8bp
pLR1687-16
Allele 1: GCTCGTTGCAATCAAGGGCACC-------CATCTCAGGTACCCCT -7bp
Allele 2: GCTCGTTGCAATCAAGGGCACCA--GCAGCATCTCAGGTACCCCT -2bp
pLR1687-21
Allele 1: GCTCGTTGCAATCAAGGGC------------TCTCAGGTACCCCT -12bp
Allele 2: GCTCGTTGCAATCAAGGGCACCATGGCAGCAgCa---------- -14bp
```

c



d

AacCas12b reagents	Tested T0 lines	Mutated T0 lines (number; percentage of total)	Biallelic mutation lines (number; percentage of total)
OsROC5-sgRNA02-L4 (pLR1809)	24	1; 4.2%	0; 0%
OsEPFL9-sgRNA02-L10 (pLR1809)	24	12; 50.0%	3; 12.5%
OsGS3-sgRNA02-L12 (pLR1809)	24	6; 25.0%	0; 0%



e

AacCas12b reagents	Tested T0 lines	Mutated T0 lines (number; percentage of total)	Biallelic mutation lines (number; percentage of total)
OsROC5-sgRNA02-L4 (pLR1811)	24	0; 0%	0; 0%
OsEPFL9-sgRNA02-L10 (pLR1811)	24	16; 66.7%	7; 29.2%
OsGS3-sgRNA02-L12 (pLR1811)	24	17; 70.8%	11; 45.8%

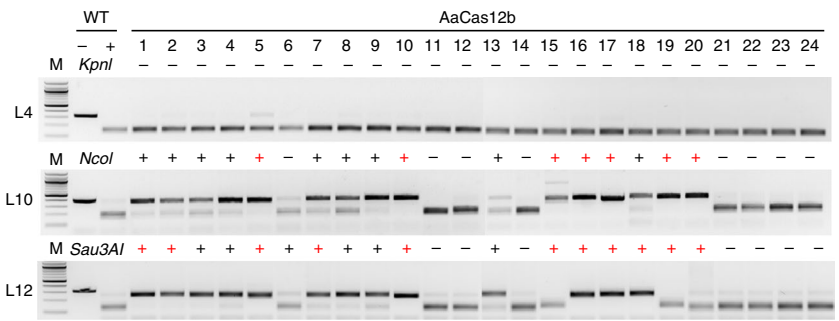


Fig. 2 | Singular and multiplexed gene editing in rice T0 lines by AacCas12b and AaCas12b. a,**b**, Summary of the genotyping results on stable transgenic T0 lines at the OsEPFL9-crRNA02 site by AacCas12b (**a**) and AaCas12b (**b**). Genotypes of five example mutants are shown for each Cas12b. The PAM sequence is in red and the target sequence is in blue. The *Nco*I enzyme site used in RFLP analysis is underlined. **c**, Illustration of the dual Pol II promoter based multiplexed Cas12b systems for AacCas12b and AaCas12b. **d**,**e**, Summary of the genotyping results on multiplexed stable transgenic T0 lines by AacCas12b (**d**) and AaCas12b (**e**) at three sites: OsROC5-sgRNA02 (L4), OsEPFL9-sgRNA02 (L10) and OsGS3-sgRNA02 (L12). RFLP analysis of independent T0 lines (shown below) and Sanger sequencing were both used for genotyping. The plus sign '+' indicates heterozygous or homozygous (highlighted in red) mutants confirmed by both methods. The data shown in **d**,**e** were combined from two independent experiments with T0 lines 1–12 generated in experiment one and T0 lines 13–24 generated in experiment two. WT, wild type.

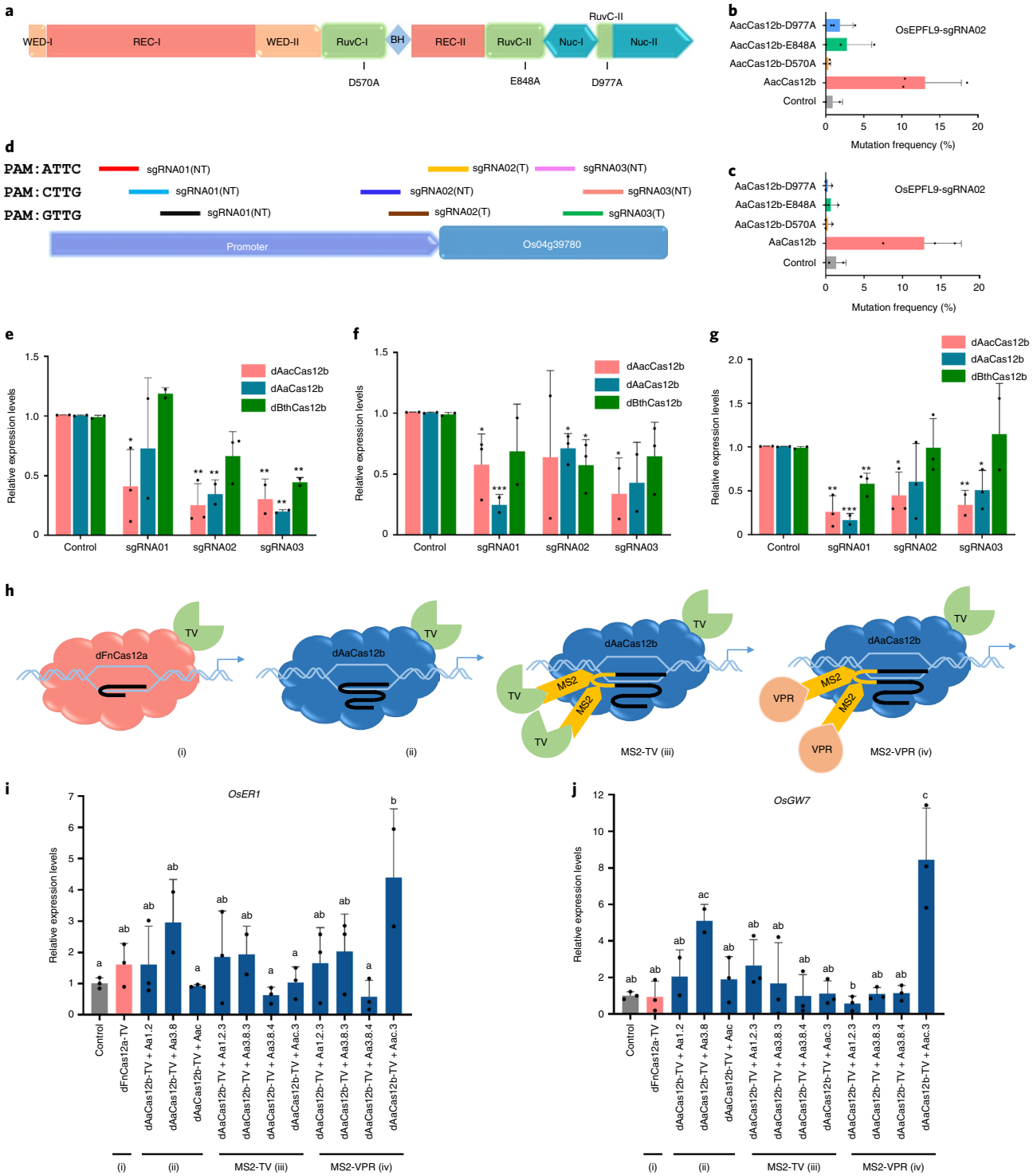


Fig. 3 | Effective CRISPR interference and CRISPR activation by dCas12b systems. a, Illustration of protein domains of AacCas12b. The three amino acid mutations used to inactivate Cas12b nuclease activity are indicated. **b**,**c**, RFLP analysis of nuclease activity for protein variants of AacCas12b (**b**) and AaCas12b (**c**) in rice protoplasts. **d**, Illustration of nine sgRNAs that direct targeted transcriptional repression at Os04g39780. Relative targeting positions and PAM sites are indicated. These sgRNAs target either the non-template strand (NT) or the template strand (T) of the DNA. **e**–**g**, qRT-PCR data showing targeted repression of Os04g39780 in rice protoplasts. dAacCas12b, dAaCas12b and dBthCas12b were compared at three different PAMs at different target positions: ATTC (**e**), CTTG (**f**) and GTTG (**g**). Statistical differences were determined using two-tailed student's *t*-tests: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, compared with the control. **h**, Schematics of four activation systems tested, including dFnCas12a-TV (i), dAaCas12b-TV with three different sgRNA scaffolds (ii), dAaCas12b-TV with four different sgRNA scaffolds containing an MS2 aptamer to recruit MS2-TV (iii) and dAaCas12b-TV with four different sgRNA scaffolds containing an MS2 aptamer to recruit MS2-VPR (iv). **i**,**j**, **j**qRT-PCR data showing targeted activation of *OsER1* (**i**) and *OsGW7* (**j**) in rice protoplasts. A total of 12 activation systems were tested. A sgRNA was used to direct each Cas12 activation system to the promoter of interest (see Supplementary Fig. 20). *OsTubulin* was used as the endogenous control gene. The gene expression level of the wild type was normalized as 1. Treatments denoted by different letters indicate significant differences between them ($P < 0.05$), as determined by Tukey's test. All data shown were mean values of three biological replicates. Error bars represent standard deviations.

efficiency at ATTG-02 site (Supplementary Fig. 19b, c). Next, we sought to develop Cas12b transcriptional activation systems based on these MS2-containing scaffolds. A potent transcriptional activator, TV[18], was fused to the C terminus of dAaCas12b. A dFnCas12a-TV fusion was also generated for comparison between Cas12a and Cas12b. Two potent activators, TV and VPR[19], were tested for MS2-based recruitment, respectively. A total of 12 transcriptional activation configurations based on four general systems were tested (Fig. 3h). Two genes, *OsER1* and *OsGW7*, were separately targeted for transcriptional activation and in each case only one sgRNA was used (Supplementary Fig. 20). Two activation systems resulted in significant transcriptional activation, while the other ten systems, including dFnCas12a-TV, failed to achieve this (Fig. 3i,j). The dAaCas12b-TV, when coupled with Aa3.8 sgRNA scaffold, resulted in three- to fivefold activation of both target genes (Fig. 3i,j). Stronger transcriptional activation (five- to eightfold) was achieved with the transcriptional system that is based on dAaCas12b-TV and Aac.3 sgRNA scaffold-mediated recruitment of MS2-VPR (Fig. 3i,j). Hence, we demonstrated a potent AaCas12b transcriptional activation system with simultaneous recruitment of TV and VPR by the dAaCas12b protein and engineered Aac.3 sgRNA, respectively.

Our work provided a demonstration for Cas12b-mediated genome editing in plants. In eukaryotic cells, we demonstrated CRISPRi with dCas12b orthologs. In addition, we developed a potent Cas12b transcriptional activation system in plants, demonstrating that Cas12b, not Cas12a, is more amendable for versatile guide RNA engineering. This feature of Cas12b may allow for future development of novel prime editing systems, as recently demonstrated with Cas9 (ref. [20]). While Cas12b nucleases often require higher temperature for optimal activity[5,6], their performance at ambient temperature, which is more relevant for plant applications, may be improved by both protein engineering and sgRNA engineering[6,17]. In conclusion, Cas12b represents another promising CRISPR system for plant genome engineering.

Methods

Construction of Gateway compatible CRISPR–Cas12b vectors. Details about construction of Gateway modular vectors for Cas12b and sgRNA are available in the Supplementary Methods. The oligos and gBlocks in this study were summarized in Supplementary Table 1.

Assembly of T-DNA expression vectors. The Cas12b T-DNA expression vectors (Supplementary Table 2) were assembled from a single Multi-site Pro LR reaction (1–5–2) with the attR1-attR2 destination vector pYPQ203 (Addgene no. 86207), an attL1-attR5 Cas12b entry clone and an attL5-attL2 crRNA expression entry clone using Gateway LR clonase II (Invitrogen). The detailed procedure is described in the Supplementary Methods.

Rice protoplast transformation and stable transformation. The Japonica cultivar Nipponbare was used in this study. Polyethylene glycol transfection of rice protoplasts was carried out according to our previously published protocol[21]. Rice stable transformation was conducted following the same procedure that we published previously[21]. Both transfection and cocultivation with *Agrobacterium* procedures were carried out at 32 °C.

Mutagenesis analysis. After 48 h of rice protoplast transfection, plant genomic DNA was extracted with the CTAB (cetyl trimethylammonium bromide) method. Nuclease-induced mutations were generally first detected and quantified by the RFLP analysis and followed by high-throughput sequencing. For high-throughput sequencing, the genomic regions flanking Cas12b target sites were PCR-amplified using barcoded primers. The PCR amplicons were sequenced by Novogene with an Illumina HiseqX platform. CRISPRMatch[22] was used to analyse the sequencing data. Both RFLP and Sanger sequencing were used to genotype individual T0 lines. Sanger sequencing results were decoded using DSDecodeM-CRISPR-GE[23]. Sanger sequencing was used to detect possible mutations at putative off-target sites, which were predicted by Cas-OFFinder[24].

RNA extraction and qRT–PCR analysis. For each sample, 400 µl (2×10^6) of rice protoplast cells transfected with 40 µg of plasmids were used for RNA extraction. The total RNA was extracted from collected protoplasts 48 h after transfection using TRIzol RNA Isolation Reagents (Thermo Fisher Scientific), and genomic DNA from RNA samples was cleaved by RNase-free DNase I (New England Biolabs) according to the manufacturer's instructions. Total RNA of 1 µg was used to synthesize the first-strand complementary DNA using the SuperScript III First-Strand Synthesis System (Thermo Fisher Scientific). qRT–PCR was performed in a CFX96 Touch Real-Time PCR Detection System (Bio-Rad) using SYBR Select Master Mix (Thermo Fisher Scientific). *OsTubulin* was used as the internal control, and fold changes were calculated by the $2^{-\Delta\Delta Ct}$ method. Three biological replicates and three technical replicates were performed for each sample. Primers used for qRT–PCR are listed in Supplementary Table 1.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The 29 Gateway compatible vectors for the CRISPR–Cas12b systems are available from Addgene: pYPQ290 (no. 129670), pYPQ291 (no. 129671), pYPQ292 (no. 129672), pYPQ290-D570A (no. 129673), pYPQ290-D977A (no. 129674), pYPQ290-E848A (no. 129675), pYPQ291-D573A (no. 129676), pYPQ291-D951A (no. 129677), pYPQ291-E827A (no. 129678), pYPQ292-D570A (no. 129679), pYPQ292-D977A (no. 129680), pYPQ292-E848A (no. 129681), pYPQ290-D570A-SRDX (no. 129682), pYPQ291-D573A-SRDX (no. 129683), pYPQ292-D570A-SRDX (no. 129684), pYPQ141-ZmUbi-RZ-Aac (no. 129685), pYPQ141-ZmUbi-RZ-Bth (no. 129686), pYPQ141-ZmUbi-RZ-Aa1.2.3 (no. 136372), pYPQ141-ZmUbi-RZ-Aa1.2 (no. 136373), pYPQ141-ZmUbi-RZ-Aa3.8.3 (no. 136374), pYPQ141-ZmUbi-RZ-Aa3.8.4 (no. 136375), pYPQ141-ZmUbi-RZ-Aa3.8 (no. 136376), pYPQ141-ZmUbi-RZ-Aac.3 (no. 136377), pYPQ141-ZmUbi-RZ-Bh (no. 136378), pYPQ239A (dFnCas12a)-TV (no. 136379), pYPQ292 (AaCas12b)-D570-TV (no. 136380), pYPQ292 (AaCas12b)-D570-TV-MS2-TV (no. 136381), pYPQ292 (AaCas12b)-D570-TV-MS2-VPR (no. 136382) and pYPQ293 (BhCas12b_v4) (no. 136383). The high-throughput sequencing data sets have been submitted to the National Center for Biotechnology information (NCBI) database under Sequence Read Archive (SRA) BioProject ID PRJNA553352.

Received: 17 July 2019; Accepted: 31 January 2020;
Published online: 09 March 2020

References

1. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
2. Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).
3. Zhang, Y., Malzahn, A., Sretenovic, S. & Qi, Y. The emerging and uncultivated potential of CRISPR technology in plant science. *Nat. Plants* **5**, 778–791 (2019).
4. Shmakov, S. et al. Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. *Mol. Cell* **60**, 385–397 (2015).
5. Teng, F. et al. Repurposing CRISPR-Cas12b for mammalian genome engineering. *Cell Discov.* **4**, 63 (2018).
6. Strecker, J. et al. Engineering of CRISPR–Cas12b for human genome editing. *Nat. Commun.* **10**, 212 (2019).
7. Yang, H., Gao, P., Rajashankar, K. R. & Patel, D. J. PAM-dependent target DNA recognition and cleavage by C2c1 CRISPR-Cas endonuclease. *Cell* **167**, 1814–1828 (2016).
8. Liu, L. et al. C2c1-sgRNA complex structure reveals RNA-guided DNA cleavage mechanism. *Mol. Cell* **65**, 310–322 (2017).
9. Wu, D., Guan, X., Zhu, Y., Ren, K. & Huang, Z. Structural basis of stringent PAM recognition by CRISPR-C2c1 in complex with sgRNA. *Cell Res.* **27**, 705–708 (2017).
10. Tang, X. et al. A CRISPR–Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nat. Plants* **3**, 17018 (2017).
11. Zhong, Z. et al. Plant genome editing using FnCpf1 and LbCpf1 nucleases at redefined and altered PAM sites. *Mol. Plant* **11**, 999–1002 (2018).
12. Jain, I. et al. Defining the seed sequence of the Cas12b CRISPR-Cas effector complex. *RNA Biol.* **16**, 413–422 (2019).
13. Paul, J. W. 3rd & Qi, Y. CRISPR/Cas9 for plant genome editing: accomplishments, problems and prospects. *Plant Cell Rep.* **35**, 1417–1427 (2016).
14. Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. Improving CRISPR–Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284 (2014).
15. Lowder, L. G. et al. A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. *Plant Physiol.* **169**, 971–985 (2015).
16. Lowder, L. G. et al. Robust transcriptional activation in plants using multiplexed CRISPR-Act2.0 and mTALE-act systems. *Mol. Plant* **11**, 245–256 (2018).
17. Teng, F. et al. Artificial sgRNAs engineered for genome editing with new Cas12b orthologs. *Cell Discov.* **5**, 23 (2019).
18. Li, Z. et al. A potent Cas9-derived gene activator for plant and mammalian cells. *Nat. Plants* **3**, 930–936 (2017).

19. Chavez, A. et al. Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* **12**, 326–328 (2015).
20. Anzalone, A. V. et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).
21. Tang, X. et al. A single transcript CRISPR-Cas9 system for efficient genome editing in plants. *Mol. Plant* **9**, 1088–1091 (2016).
22. You, Q. et al. CRISPRMatch: An automatic calculation and visualization tool for high-throughput CRISPR genome-editing data analysis. *Int J. Biol. Sci.* **14**, 858–862 (2018).
23. Liu, W. et al. DSDecode: A web-based tool for decoding of sequencing chromatograms for genotyping of targeted mutations. *Mol. Plant* **8**, 1431–1433 (2015).
24. Bae, S., Park, J. & Kim, J. S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473–1475 (2014).

Acknowledgements

This work was supported by University of Maryland startup funds, the National Science Foundation Plant Genome Research Program grant (award no. IOS-1758745), the Biotechnology Risk Assessment Grant Program competitive grant (award no. 2018-33522-28789) from the US Department of Agriculture, Foundation for Food and Agriculture Research grant (award no. 593603) and Syngenta Biotechnology to Y.Q. It was also supported by the National Transgenic Major Project (award nos. 2019ZX08010003-001-002 and 2018ZX08020-003), the National Science Foundation of China (award no. 31771486), the Sichuan Youth Science and Technology Foundation (award no. 2017JQ0005) and the Science Strength Promotion Program of UESTC to Yong Z. M.M was supported by a scholarship from China Scholarship Council. A.M was supported by a scholarship from Cosmos Club Foundation.

Author contributions

Y.Q. and Yong Z. designed the experiments. M.M, Yingxiao Z. and C.P. generated all the constructs. M.M., Q.R., Y.H., S.L., Z.Z., J.W. and X.Z. performed the transient assays in protoplasts. Q.R. and Y.H. prepared samples for high-throughput sequencing. M.M. generated stable transgenic rice and analysed the plants. C.P. conducted transcriptional repression and activation assays. Y.Q., Yong Z., A.M. and J.W. wrote the paper with input from other authors. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41477-020-0614-6.

Correspondence and requests for materials should be addressed to Y.Z. or Y.Q.

Peer review information *Nature Plants* thanks Sang Gyu Kim and Huawei Zhang and the other, anonymous, reviewer for their contribution to the peer review of this work.

nature research

Corresponding author(s): Yiping Qi

Last updated by author(s): 2019/01/22

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data collection details are available in Methods.
Data analysis	CRISPRMatch, DSDecodeM-CRISPR-GE, Cas-OFFinder

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All 29 Cas12b Gateway entry clones for plant genome engineering have submitted to Addgene for public dissemination; The high-throughput sequencing data sets have been submitted to the National Center for Biotechnology information (NCBI) database under Sequence Read Archive (SRA) BioProject ID: PRJNA553352.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were predetermined based on the convention in the field of plant genome editing. For example, at least two independent sites were assessed for each PAM requirement. For each construct, three biological replicates were used for assessing editing activity. For stable transformation, 12 or more individual transgenic plants were evaluated for each construct in each experiment.
Data exclusions	No data was excluded.
Replication	The experiments were replicated independently at least once.
Randomization	Not relevant in our experiments as we do not expect any possible biases without doing so.
Blinding	Not relevant in our experiments as we do not expect any possible biases without doing so.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging