

# Expanding the scope of plant genome engineering with Cas12a orthologs and highly multiplexable editing systems

Yingxiao Zhang<sup>1,9</sup>, Qiurong Ren<sup>2,9</sup>, Xu Tang<sup>2,9</sup>, Shishi Liu<sup>2</sup>, Aimee A. Malzahn<sup>1</sup>, Jianping Zhou<sup>2</sup>, Jiaheng Wang<sup>2</sup>, Desuo Yin<sup>1,3</sup>, Changtian Pan<sup>1</sup>, Mingzhu Yuan<sup>2</sup>, Lan Huang<sup>2</sup>, Han Yang<sup>2</sup>, Yuxin Zhao<sup>2</sup>, Qing Fang<sup>2</sup>, Xuelian Zheng<sup>2</sup>, Li Tian<sup>2</sup>, Yanhao Cheng<sup>1,4</sup>, Ysa Le<sup>1</sup>, Bailey McCoy<sup>1</sup>, Lidiya Franklin<sup>1</sup>, Jeremy D. Selengut<sup>5</sup>, Stephen M. Mount<sup>6</sup>, Qiudeng Que<sup>7</sup>, Yong Zhang<sup>2</sup>✉ & Yiping Qi<sup>1,8</sup>✉

CRISPR-Cas12a is a promising genome editing system for targeting AT-rich genomic regions. Comprehensive genome engineering requires simultaneous targeting of multiple genes at defined locations. Here, to expand the targeting scope of Cas12a, we screen nine Cas12a orthologs that have not been demonstrated in plants, and identify six, ErCas12a, Lb5Cas12a, BsCas12a, Mb2Cas12a, TsCas12a and MbCas12a, that possess high editing activity in rice. Among them, Mb2Cas12a stands out with high editing efficiency and tolerance to low temperature. An engineered Mb2Cas12a-RVRR variant enables editing with more relaxed PAM requirements in rice, yielding two times higher genome coverage than the wild type SpCas9. To enable large-scale genome engineering, we compare 12 multiplexed Cas12a systems and identify a potent system that exhibits nearly 100% biallelic editing efficiency with the ability to target as many as 16 sites in rice. This is the highest level of multiplex edits in plants to date using Cas12a. Two compact single transcript unit CRISPR-Cas12a interference systems are also developed for multi-gene repression in rice and *Arabidopsis*. This study greatly expands the targeting scope of Cas12a for crop genome engineering.

<sup>1</sup>Department of Plant Science and Landscape Architecture, University of Maryland, College Park, MD, USA. <sup>2</sup>Department of Biotechnology, School of Life Science and Technology, Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu, China. <sup>3</sup>Food Crop Institute, Hubei Academy of Agricultural Sciences, Wuhan, Hubei, China. <sup>4</sup>College of Agriculture, Nanjing Agricultural University, Nanjing, Jiangsu, China. <sup>5</sup>Center for Bioinformatics and Computational Biology, University of Maryland, College Park, MD, USA. <sup>6</sup>Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD, USA. <sup>7</sup>Syngenta, Research Triangle Park, NC, USA. <sup>8</sup>Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, MD, USA. <sup>9</sup>These authors contributed equally: Yingxiao Zhang, Qiurong Ren, Xu Tang. ✉email: [zhangyong916@uestic.edu.cn](mailto:zhangyong916@uestic.edu.cn); [yiping@umd.edu](mailto:yiping@umd.edu)

CRISPR-Cas12a (formerly Cpf1) is a class II type V endonuclease that prefers a thymine-rich protospacer adjacent motif (PAM)<sup>1</sup>, and is the second most commonly used CRISPR system for genome editing<sup>2–4</sup>. Cas12a has shown higher targeting specificity than Cas9 in mammalian cells and plants<sup>5–7</sup>. Unlike Cas9, Cas12a generates staggered ends and larger deletions, making it a suitable nuclease for gene knockout. Moreover, Cas12a only requires a short CRISPR RNA (crRNA) for each target and possesses RNase activity for crRNA array processing, making it an excellent platform for multiplexed editing<sup>8,9</sup>. Cas12a has been used to generate targeted mutations as well as achieve transcriptional regulation in microorganisms<sup>10–12</sup>. Cas12a can also efficiently generate edits in some important industrial *Streptomyces* strains that cannot be edited using SpCas9 due to toxicity<sup>12</sup>. In mammalian systems, Cas12a has been widely used and engineered for genetic manipulation<sup>1,13–15</sup>. Three Cas12a orthologs, LbCas12a, AsCas12a, and FnCas12a, have been used in plants. LbCas12a is most popular due to its high editing activity in rice<sup>9,16–22</sup>, maize<sup>23,24</sup>, *Arabidopsis*<sup>24–26</sup>, tomato<sup>25</sup>, *Nicotiana benthamiana*<sup>25</sup>, lettuce<sup>27</sup>, cotton<sup>28</sup>, and citrus<sup>29</sup>. FnCas12a also mediates efficient genome editing in plants<sup>9,17,20,30</sup>. While most Cas12a studies in plants pursued targeted mutagenesis by non-homologous end joining (NHEJ) DNA repair, precise genome editing based on homology-directed repair (HDR) has also been demonstrated<sup>20,22,31–33</sup>. Furthermore, Cas12a has been repurposed for base editing<sup>15,34</sup> and transcriptional regulation<sup>15,16,35–37</sup>.

Plant genome editing is an exciting technology to accelerate crop breeding. The versatility and simplicity of the CRISPR-Cas12a system allow simultaneous modification of many genes in elite cultivars, that may be difficult or impractical to achieve with conventional breeding and genetic engineering methods. However, the VTTV (V = A, C, and G) PAM requirement of Cas12a limits its targeting scope. Many attempts have been made to engineer Cas12a nucleases<sup>13–15,38</sup> or explore novel Cas12a<sup>1,14,39–41</sup> with relaxed PAMs in vitro and in mammalian cells. In plants, although FnCas12a has been demonstrated to target some VTTV PAMs<sup>17,30</sup> and engineered LbCas12a can target altered PAMs<sup>17,42</sup>, further improvement is required to edit more relaxed PAMs. The second limitation is that current Cas12a orthologs in use are not effective at lower temperatures<sup>24,43</sup>, though this could partially be addressed by protein engineering<sup>15,26</sup>. Moreover, a multiplexing system with high efficiency and capacity is required to enable versatile genome engineering in plants. Although many multiplexing strategies have been developed in plants using Cas9 and Cas12a (Supplementary Tables 1 and 2), a system that can achieve highly efficient biallelic editing at multiple targets has not been demonstrated. With low biallelic editing efficiency, it is laborious or impractical to obtain multigene knockouts in crops through subsequent breeding. To address these issues, we comprehensively assess nine Cas12a orthologs and engineered their variants in rice for efficient genome editing with expanded targeting ranges. We also compared 12 multiplexed Cas12a systems to identify the most potent one for large-scale plant genome engineering. Collectively, we establish a Cas12a toolbox for efficient multiplexed genome engineering with improved target accessibility and scalability.

## Results

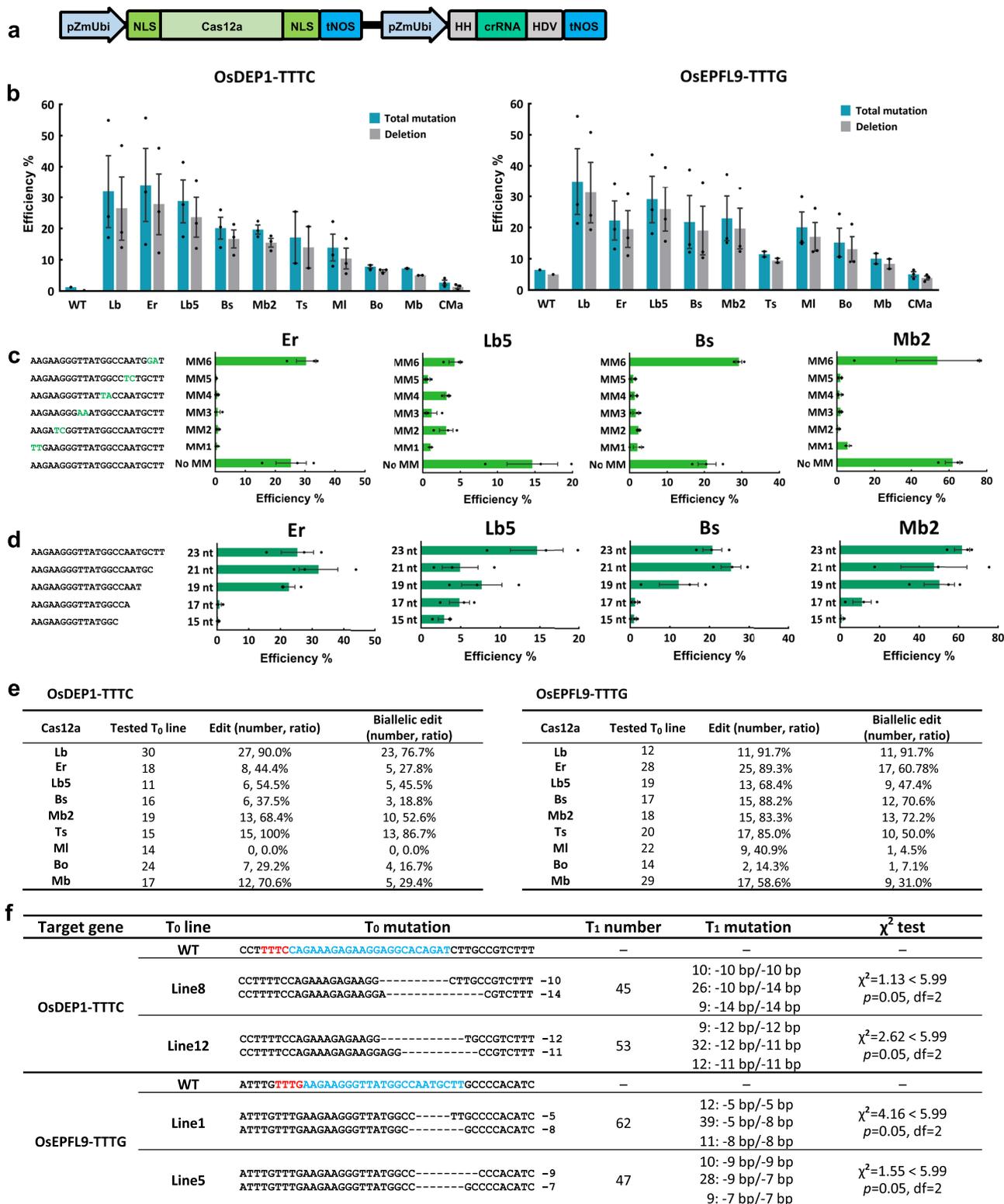
**Genome editing using nine Cas12a orthologs in rice cells.** To improve Cas12a genome editing in plants, we chose nine Cas12a orthologs that have not been demonstrated in plants (Supplementary Fig. 1). We included eight Cas12a orthologs that showed a preference for TTV PAMs in vitro, including Lb5Cas12a, BsCas12a, Mb2Cas12a, TsCas12a, MlCas12a, BoCas12a, MbCas12a, and CMacas12a<sup>1,39</sup>. In addition, we included ErCas12a<sup>44</sup>, which is also known as MAD7 (Inscripta, Inc.). The editing activity of these

Cas12a orthologs, along with the control LbCas12a, was evaluated in rice protoplasts at two target sites, OsDEP1-TTTC and OsEPFL9-TTTG, using a dual RNA Polymerase II (Pol II) promoter expression system<sup>16</sup> (Fig. 1a). Editing efficiencies were assessed by high-throughput amplicon sequencing using the Illumina HiSeq platform. All Cas12a orthologs, except CMacas12a, resulted in targeted mutation frequencies of roughly 10% or greater at both sites (Fig. 1b and Supplementary Data 1). Four Cas12a orthologs (Er, Lb5, Bs, and Mb2) showed comparable editing frequencies to LbCas12a. Targeted mutations were predominately deletions distal from PAMs, centered around the predicted cleavage site<sup>1</sup> (Fig. 1b and Supplementary Figs. 2 and 3). When mismatched protospacers were used, all four Cas12a orthologs could only tolerate mismatches at the last three base pairs distal from the PAM, indicating high targeting specificity (Fig. 1c, Supplementary Fig. 4, and Supplementary Data 1). Protospacers of 19–23 nt could generate high editing frequencies for ErCas12a and Mb2Cas12a, while longer protospacers were required for Lb5Cas12a and BsCas12a to achieve robust activity (Fig. 1d, Supplementary Fig. 4, and Supplementary Data 1).

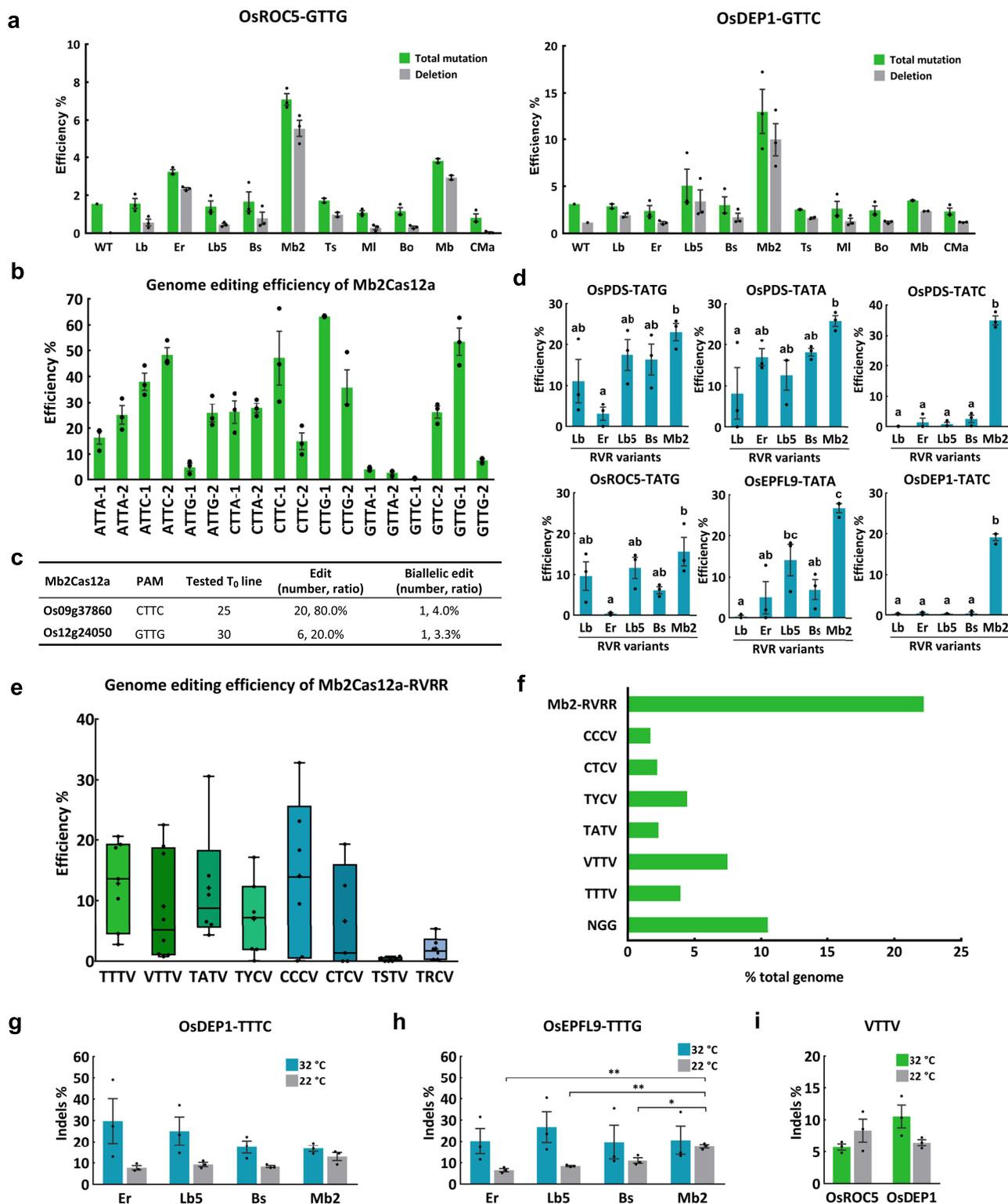
**Genome editing using eight Cas12a orthologs in stable transgenic rice.** We next assessed these eight Cas12a orthologs (except CMacas12a) in stable transgenic rice lines at the same two sites. All eight Cas12a orthologs could generate biallelic mutants in rice (Fig. 1e and Supplementary Figs. 5, 6) and most of them (e.g., Er, Lb5, Bs, Mb2, Ts, and Mb) showed medium to high editing frequencies (Fig. 1e). At the OsDEP1-TTTC site, TsCas12a showed 100% editing frequency, which was even higher than LbCas12a; three other Cas12a orthologs (Lb5, Mb2, and Mb) also had high editing frequencies from 54.5 to 70.6% (Fig. 1e). At the OsEPFL9-TTTG site, four Cas12a orthologs (Er, Bs, Mb2, and Ts) showed comparable editing frequencies (83.3–89.5%) to LbCas12a (91.7%) (Fig. 1e). Together, these data suggest the tested Cas12a orthologs can induce medium to high-frequency genome editing in rice except for MlCas12a and BoCas12a. To assess the inheritance of targeted mutations, we examined Mb2Cas12a  $T_1$  populations derived from two  $T_0$  lines for each target site. In all cases, mutations were stably inherited with the expected 1:2:1 Mendelian segregation ratio (Fig. 1f).

**Mb2Cas12a can efficiently target VTTV PAMs.** To assess whether the tested Cas12a orthologs can target shorter PAMs, we measured their editing activity at two relaxed VTTV PAM sites, OsROC5-GTTG and OsDEP1-GTTC. LbCas12a and most of the Cas12a orthologs showed low efficient or no editing (Fig. 2a and Supplementary Data 1). However, Mb2Cas12a displayed editing efficiencies of ~10% at both target sites (Fig. 2a and Supplementary Data 1) and resulted in largely 3–12 bp deletions (Supplementary Fig. 7). We next tested Mb2Cas12a at 18 target sites with two sites for each possible VTTV combination (Supplementary Fig. 8). Mb2Cas12a could efficiently edit 13 out of 18 target sites with mutation frequencies of approximately 15% or greater (Fig. 2b and Supplementary Data 1). Among these 13 sites, ATTA-1 and ATTC-2, which could not be edited by FnCas12a<sup>17</sup>, were edited by Mb2Cas12a with efficiencies of 16.3% and 48.3%, respectively. Moreover, Mb2Cas12a could edit GTTA and GTTC PAM sites, which were not editable by FnCas12a<sup>17</sup>. Further, we found Mb2Cas12a could reliably generate stable knockouts of two target genes at VTTV PAM sites (Fig. 2c and Supplementary Fig. 9). Taken together, Mb2Cas12a can reliably target almost all NTTV PAMs in rice.

**The RVR variant of Mb2Cas12a can efficiently target TATV PAMs.** Previously, RVR variants of AsCas12a and LbCas12a were



**Fig. 1 Targeted mutagenesis by Cas12a orthologs at canonical TTTV PAM sites in rice.** **a** A schematic diagram of the Cas12a expression system for single site targeting. pZmUbi maize ubiquitin promoter, NLS nuclear localization signal, tNOS nopaline synthase (NOS) terminator, HH hammerhead ribozyme, HDV hepatitis delta virus ribozyme. **b** Targeted mutagenesis efficiencies (percentage) of ten Cas12a orthologs at two TTTV PAM sites in rice protoplasts. WT, protoplasts transformed with water. **c** Targeting specificity of four Cas12a orthologs measured with mismatched crRNAs at the OsEPFL9-TTTG site in rice protoplasts. Mismatched nucleotides in crRNAs are highlighted in green. **d** Protospacer length requirements of four Cas12a orthologs at the OsEPFL9-TTTG site in rice protoplasts. **e** Summary of editing and biallelic editing efficiencies of Cas12a orthologs at two TTTV PAM sites in transgenic rice T<sub>0</sub> lines. **f** Editing analysis of transgenic rice T<sub>1</sub> lines.  $\chi^2$  test results indicate that T<sub>1</sub> populations from each tested T<sub>0</sub> line segregated at a 1:2:1 ratio when  $\alpha = 0.05$ . Data in “b–d” are presented as mean values  $\pm$  SEM.  $n = 3$  biologically independent samples. Source data are provided as a Source Data file.



engineered to target TATV PAMs in mammalian cells<sup>13</sup>. We showed that LbCas12a-RVR could only target TATG PAMs while FnCas12a-RVR had no activity at TATV PAMs in rice<sup>17</sup>. To seek Cas12a variants that can truly target all TATV PAMs, we engineered RVR variants of ErCas12a, Lb5Cas12a, BsCas12a, and Mb2Cas12a (Supplementary Fig. 1b), and tested them in rice protoplasts (Supplementary Fig. 10 and Supplementary Data 1). LbCas12b-RVR<sup>17</sup> was included as a control. RVR variants of Lb5Cas12a, BsCas12a, and Mb2Cas12a showed robust editing at the two TATG PAM sites, as with LbCas12a-RVR (Fig. 2d). All

four RVR variants could edit both TATA PAM sites with higher average editing efficiencies than LbCas12a-RVR (Fig. 2d). However, only Mb2Cas12a-RVR could edit the two TATC PAM sites (Fig. 2d). Impressively, Mb2Cas12a-RVR displayed high editing frequencies at all TATV PAM sites, ranging from 15 to 35%, which outperformed all other tested RVR variants.

**The RVRR variant of Mb2Cas12a has a broad target range.** To further broaden the target range of Mb2Cas12a, we engineered the RVRR variant by introducing the K625R mutation into its

**Fig. 2 Mb2Cas12a and engineered Cas12a variants broaden the genome editing scope in rice.** **a** Targeted mutagenesis efficiencies (percentage) of ten Cas12a orthologs at two VTTV PAM sites in rice protoplasts. WT, protoplasts transformed with water. **b** Targeted mutagenesis efficiencies (percentage) of Mb2Cas12a at 18 VTTV PAM sites in rice protoplasts. **c** Summary of editing and biallelic editing efficiencies by Mb2Cas12a at two VTTV PAM sites in transgenic rice  $T_0$  lines. **d** Comparison of editing efficiencies (percentage) of five Cas12a RVR variants at six TATV PAM sites in rice protoplasts. Treatments with the same letter are not significantly different when  $\alpha = 0.05$  by Tukey's Honest Significant Difference test (two-sided). **e** A box plot showing targeted mutagenesis efficiencies (percentage) of Mb2Cas12a-RVRR variant at 51 canonical and altered PAM sites in rice protoplasts. V = A, C, and G. Y = T and C. S = C and G. R = A and G. Each data point represents the mean of three biological replicates ( $n = 3$ ). The middle line in each box denotes the median. The upper and lower bounds of each box denote the first quartile (25%) and the third quartile (75%), respectively. The top and bottom of whiskers denote the maxima and minima, respectively. **f** Genome-wide analysis of targetable PAM sites by Mb2Cas12a-RVRR variant in rice (*Oryza sativa*). **g** and **h**, Insertion and deletion (Indel) frequencies (percentage) of four Cas12a orthologs at two TTTV PAM sites at 32 and 22 °C in rice protoplasts. One asterisk ( $p < 0.05$ ) and two asterisks ( $p < 0.01$ ) indicate significant differences between two treatments using two-sided Student's *t*-test.  $p = 0.000583$  (Er vs. Mb2);  $p = 0.000456$  (Lb5 vs. Mb2);  $p = 0.012092$  (Bs vs. Mb2). **i** Indel frequencies (percentage) of Mb2Cas12a at two VTTV PAM sites at 32 and 22 °C in rice protoplasts. Data in **a**, **b**, **d**, **g**–**i** are presented as mean values  $\pm$  SEM.  $n = 3$  biologically independent samples. Source data are provided as a Source Data file.

RVR variant (Supplementary Fig. 11). We tested its editing activity at 51 targeting sites with canonical and altered PAMs in rice protoplasts (Supplementary Data 1). Mb2Cas12a-RVRR showed editing activities at TTTV, VTTV, TATV, TYCV, CCCV, and CTCV PAM sites (Y = T and C) (Fig. 2e). The variant however showed low or no editing activity at TSTV and TRCV PAM sites (S = C and G, R = A and G) (Fig. 2e). The RVRR variant of Mb2Cas12a significantly broadens the target range by Cas12a.

### Mb2Cas12a greatly expands the scope of crop genome editing.

To assess the targeting scope of Mb2Cas12a and its variants, we conducted genome-wide PAM analysis in rice (*Oryza sativa*). The analysis showed the canonical Cas12a's TTTV PAM sites only cover 4.0% of the genomes (Fig. 2f). Additional 7.5% and 2.3% genomic area become accessible with VTTV PAMs and TATV PAMs by using Mb2Cas12a and Mb2Cas12a-RVR, respectively (Fig. 2f). Remarkably, all the sites targetable by Mb2Cas12a-RVRR yield 22.2% coverage of the total genome, which is much higher than that of SpCas9 with an NGG PAM requirement (10.5%) (Fig. 2f).

**Mb2Cas12a is tolerant of low temperatures.** Cas12a generally has reduced editing activity at lower temperatures<sup>24,43</sup>. We compared the editing activity of four Cas12a orthologs at 32 and 22 °C in rice protoplasts (Supplementary Data 1). At the two TTTV PAM sites, Mb2Cas12a activity was less affected by the low temperature and appeared to have higher Insertion and deletion (Indel) frequencies than ErCas12a, Lb5Cas12a, and BsCas12a (Fig. 2g, h). Editing activities were also detected at 22 °C with Mb2Cas12a at two VTTV PAM sites (Fig. 2i), further supporting Mb2Cas12a is tolerant of low temperature. Next, we tested two Mb2Cas12a variants (v1 and v2) harboring equivalent mutations of enhanced AsCas12a (enAsCas12a)<sup>15</sup>, and found Mb2Cas12a-v2 had improved activity at the OsEPFL9-TTTG PAM site at 22 °C (Supplementary Fig. 12). These data suggest Mb2Cas12a is less temperature sensitive and its activity can be further improved with protein engineering.

### Comparison of ten multiplex Cas12a genome editing systems.

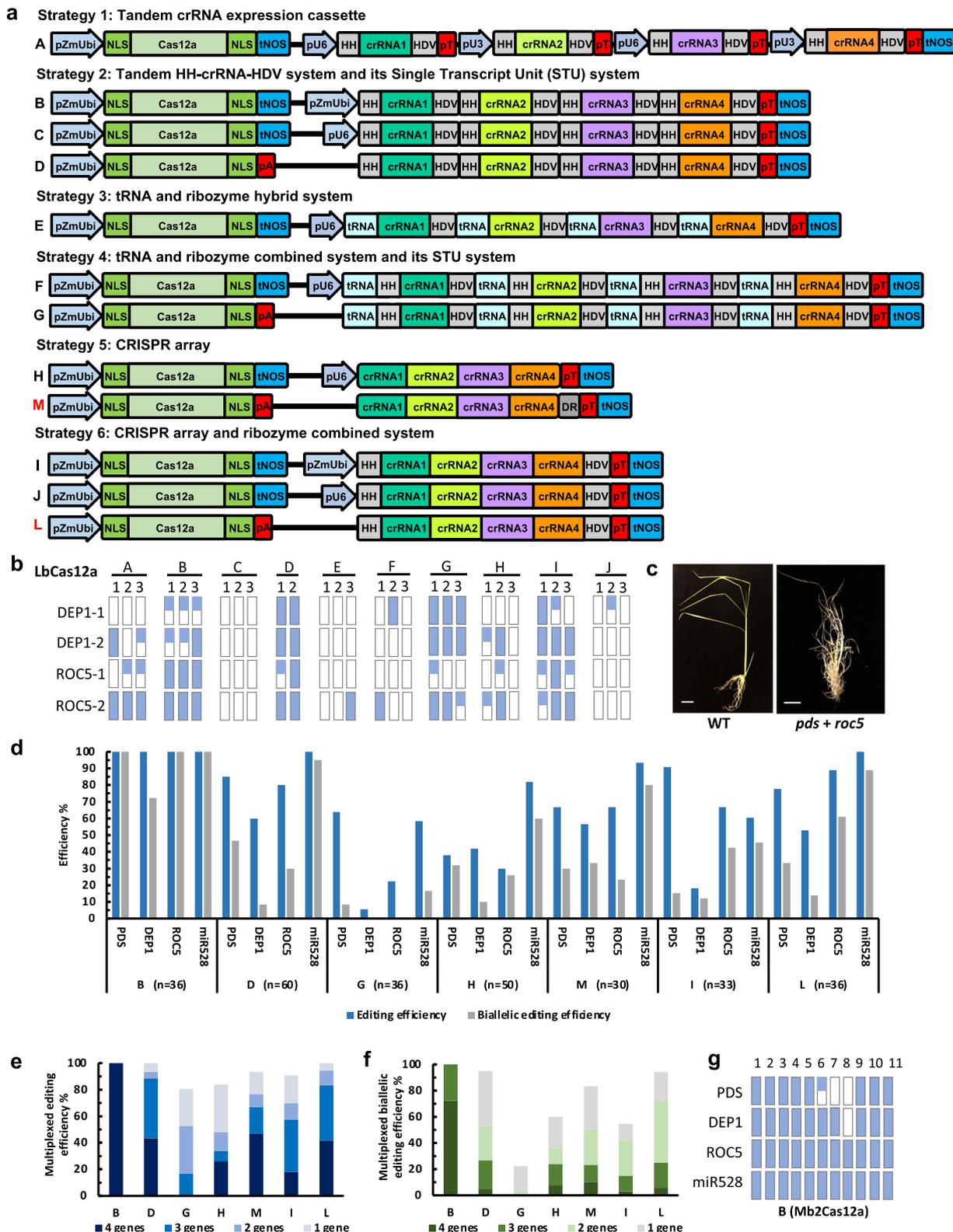
To further expand the genome engineering scope with targeting multiple sites simultaneously, we sought to develop efficient multiplexed Cas12a systems. We first compared ten multiplexed systems (A through J) that can be grouped into six strategies (Fig. 3a). Strategy 1 utilizes tandem crRNA expression cassettes where each crRNA has its own promoter, either OsU6 or OsU3. The crRNAs are processed by the hammer head (HH) and hepatitis delta virus (HDV) ribozymes<sup>16</sup>. Strategy 2 compares the tandem

HH-crRNA-HDV system under three expression conditions: by a Pol II promoter (pZmUbi), by a Pol III promoter (OsU6) or by a single transcript unit (STU) system. Strategy 3 uses a OsU6 to drive a tRNA-crRNA-HDV array, taking advantage of the tRNA's promoter activity<sup>45</sup>. Given tRNA processing may leave a few extra nucleotides at the 5' end<sup>45</sup>, Strategy 4 explores the use of a tRNA-HH-crRNA-HDV array to ensure more precise processing. Strategy 5 is the OsU6-driven CRISPR array system that has been previously explored<sup>9</sup>. Strategy 6 employs HH and HDV ribozymes to flank the entire crRNA array. Both pZmUbi and OsU6 promoters are compared in this strategy.

These ten systems were compared in stable transgenic rice lines using LbCas12a and four multiplexed crRNAs to simultaneously target two genes, *OsDEP1* and *OsROC5*. About three independent  $T_0$  lines were genotyped at each target site for each system. Based on the results, we could roughly rank these multiplex systems from high activity to low activity: System B and D > System G and I > System A > System H > System C, E, F, and J (Fig. 3b). The best performing multiplex systems (B and D) used a strong Pol II promoter to express an HH-crRNA-HDV array and resulted in high frequency biallelic mutations in  $T_0$  lines (Fig. 3b and Supplementary Fig. 13). By contrast, the least performing multiplex systems (C, E, F, and J) all used the Pol III promoter OsU6 and at most resulted in mutations at one site (Fig. 3b and Supplementary Fig. 13). These results suggest Pol III promoter-based systems are not suitable for expressing multiple crRNAs, even though they were previously used in humans<sup>8</sup> and plants<sup>9</sup>; this may be because Pol III promoters are evolved to transcribe tRNA and other small RNAs in nature and not suitable for producing long transcripts.

### Comparison of seven refined multiplex Cas12a genome editing systems.

We next focused on seven refined multiplexed LbCas12a systems that primarily used Pol II promoters: B, D, G, H, M, I, and L (Fig. 3a). We included two additional STU systems, M and L, with M using an extra direct repeat (DR) at the end of the crRNA array and L using HH and HDV to flank the crRNA array. This time, we used four crRNAs to target four different genes. For the seven systems, we generated 36, 60, 36, 50, 30, 33, and 36  $T_0$  lines, respectively. While biallelic editing at *OsPDS* and *OsROC5* could be detected based on phenotype (Fig. 3c), we relied on restriction fragment length polymorphism (RFLP) assays and Sanger sequencing to genotype the edits (Fig. 3d and Supplementary Figs. 14–20). We first calculated editing efficiencies at each target site for all the multiplex systems. System B stood out as the most efficient multiplex system, resulting in nearly 100% biallelic mutagenesis at all target sites (Fig. 3d and Supplementary Fig. 14). More impressively, all  $T_0$  lines carried biallelic mutations at three target genes: *OsPDS*, *OsROC5*, and *OsmiR528*. The high editing efficiency of system B was further confirmed with amplicon sequencing using the Hi-



TOM platform<sup>46</sup> (Supplementary Figs. 21 and 22). The second tier of high-efficiency multiplex systems includes systems D, M, and L, where editing frequencies were over 50% for every target site. The remaining three systems, I, H, and G, had relatively low multiplex editing frequencies.

To gain further insights about the multiplexing ability of these systems, we calculated their editing frequencies when editing defined numbers of genes (4, 3, 2, and 1) simultaneously. System B

again stood out since it achieved the highest multiplex editing when targeting four genes (Fig. 3e). The second-tier systems, D, M, and L, were clustered together with 40–50% quadruple editing frequencies (Fig. 3e). A similar trend has been observed when we calculated the biallelic editing efficiencies (Fig. 3f). We further applied the highly efficient multiplex system B to Mb2Cas12a and found 8 out of 11 *T<sub>0</sub>* lines carried biallelic edits at all four target sites (Fig. 3g and Supplementary Fig. 23). We then used Mb2Cas12a-

**Fig. 3 Comparison of 12 multiplex Cas12a genome editing systems in rice.** **a** Schematics of 12 multiplexed Cas12a systems, which are classified into six different strategies. The schematics are based on multiplexing four crRNAs. pZmUbi maize ubiquitin promoter, NLS nuclear localization signal, tNOS nopaline synthase (NOS) terminator, pU6 rice U6 promoter, pU3 rice U3 promoter, HH hammerhead ribozyme, HDV hepatitis delta virus ribozyme, pT polyT, pA polyA, DR direct repeat. **b** Comparison of ten multiplexed Cas12a editing systems (A–J) in stable transgenic rice lines, by using LbCas12a to target *OsDEP1* and *OsROCS5* with four crRNAs. Approximately three independent  $T_0$  lines were genotyped at each target site for wild type (denoted as an empty rectangle), monoallelic mutant (denoted as a half-filled rectangle) and biallelic mutant (denoted as a fully filled rectangle). **c** An example of a quadruple mutant showing loss-of-function phenotypes for *OsPDS* and *OsROCS5*, compared to the wild type (WT). Additional quadruple mutants with similar phenotypes were identified. Size bar, 4 cm. **d** Comparison of seven multiplexed editing systems (B, D, G, H, M, I, and L) in stable transgenic rice lines, using LbCas12a to target four genes (*OsPDS*, *OsDEP1*, *OsROCS5*, and *OsmiR528*) with four crRNAs. The number of transgenic  $T_0$  plants assayed are also indicated by “n”. **e** Further analysis of the data in “d” for editing efficiencies of each multiplexed LbCas12a system when editing certain defined numbers of genes (4, 3, 2, and 1) simultaneously. **f** Further analysis of the data in “d” for biallelic editing efficiencies of each multiplexed LbCas12a system when editing certain defined numbers of genes (4, 3, 2, and 1) simultaneously. **g** Multiplexed editing of four genes with system B by Mb2Cas12a. Genotypes at each target site in 11  $T_0$  lines are shown with wild type denoted as an empty rectangle, monoallelic mutant denoted as a half-filled rectangle and biallelic mutant denoted as a fully filled rectangle. Source data are provided as a Source Data file.

RVRR coupled with multiplex system B to target seven sites with canonical and altered PAMs simultaneously. All sites were edited with efficiencies ranging from 5.1 to 31.1% in rice protoplasts (Supplementary Fig. 24 and Supplementary Data 1). Taken together, we identified system B, a dual Pol II promoter and tandem HH-crRNA-HDV system, as the most efficient multiplex Cas12a system (Fig. 3a).

To investigate whether multiplexed genome editing by Mb2Cas12a can cause off-target mutagenesis, we amplified and sequenced 27 potential off-target sites with four or fewer mismatches or bulges within the protospacers in two transgenic  $T_0$  lines. No off-target mutations were detected (Supplementary Table 3), suggesting high specificity of Mb2Cas12a during multiplexed genome editing.

**Multitrait engineering within one generation.** We next applied the multiplex system B for simultaneous editing of five trait genes to enhance yield and blight disease resistance in rice. We generated a Cas12a vector multiplexing six crRNAs to edit the coding sequences of *OsGS3* (*GRAIN SIZE3*)<sup>47</sup> and *OsGW2* (*GRAIN WIDTH and WEIGHT2*)<sup>48</sup>, as well as the promoters of three *SWEET* genes, *SWEET11*, *SWEET13*, and *SWEET14*. *OsGS3* and *OsGW2* are negatively correlated with grain size, and grain width and weight, respectively. Transcription activator-like effectors (TALEs) from *Xanthomonas oryzae* pv. *oryzae* can bind to the *SWEET* promoter region, causing bacterial leaf blight<sup>49</sup>. Mb2Cas12a was chosen since one target site contains a VTTV PAM. Four crRNAs were designed to destroy all the TALE binding sites within the *SWEET* promoters, including PthXo1, TalC, TalF, AvrXa7, and PthXo3<sup>49,50</sup> (Fig. 4a). High biallelic editing efficiencies at all six target sites in 8 out of 10 individual  $T_0$  lines were observed (Fig. 4a and Supplementary Fig. 25). All the targeted TAL effector binding sites were mutated by large deletions in all these lines, except for TalC. Expression of the *SWEET* genes was not compromised in these promoter-edited lines (Supplementary Fig. 26), indicating that the promoter function of the *SWEET* genes was not disrupted. These results reinforced the practical application of this promoter editing strategy as demonstrated with TAL effector nuclease (TALEN)<sup>51</sup> and Cas9<sup>49</sup>. Simultaneous editing of *OsGS3* and *OsGW2* was also achieved in these lines (Fig. 4a), which is expected to contribute to increased yield<sup>52</sup>. Off-target analysis of two multiplex edited  $T_0$  lines showed no edits at potential off-target sites (Supplementary Table 3). Together, we demonstrated the use of a multiplexed Mb2Cas12a system for simultaneous biallelic editing of different trait genes in rice within one generation.

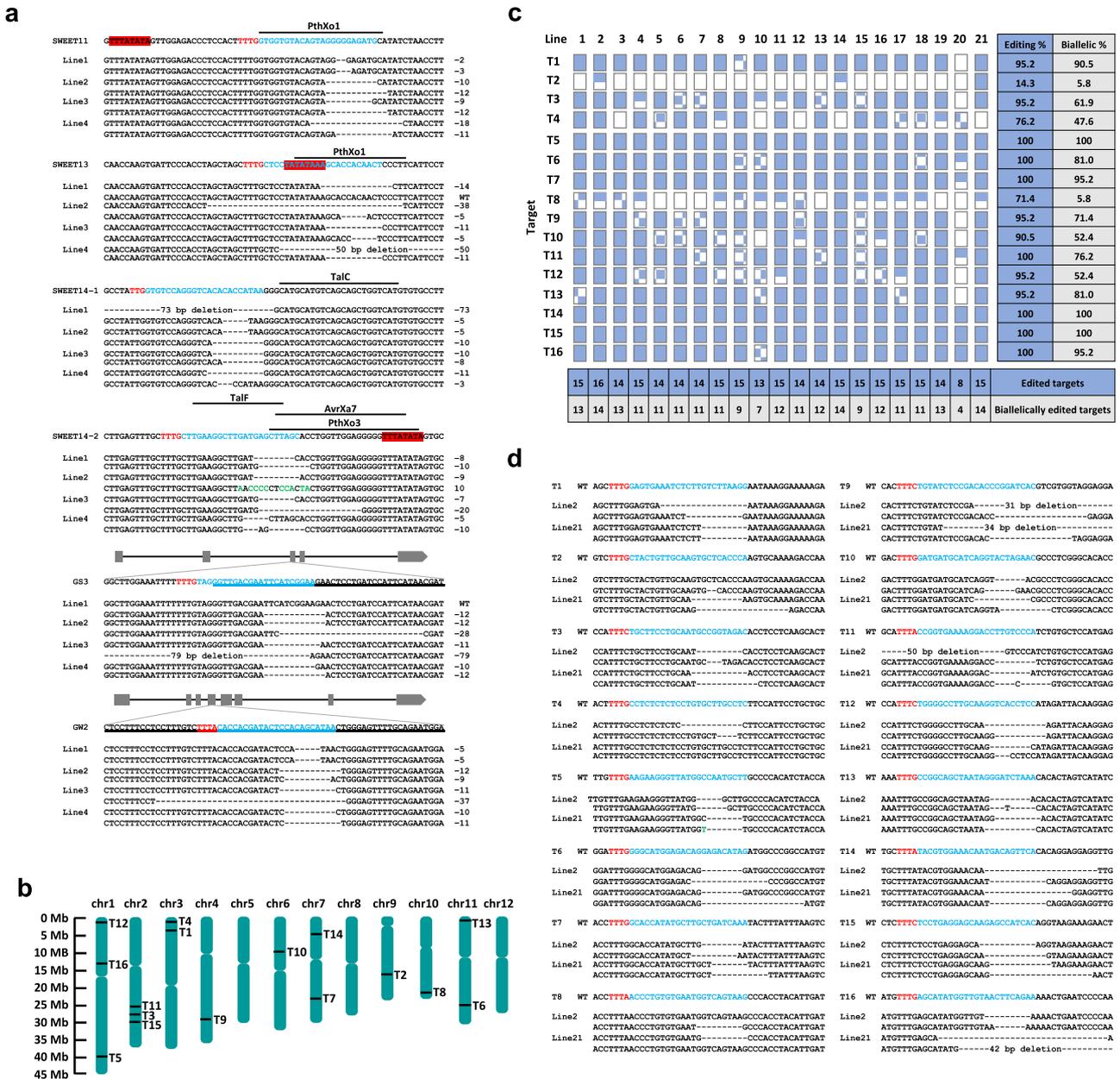
**Large-scale biallelic genome editing.** Since the multiplex system B showed nearly 100% biallelic editing frequencies at four and six target sites (Figs. 3 and 4a), we would like to see how it performs

when scaled up. To this end, we generated a single LbCas12a T-DNA vector to simultaneously target 16 genomic sites across nine chromosomes in rice (Fig. 4b). We assessed editing at all sites in 21  $T_0$  lines with RFLP assays and the Hi-TOM sequencing platform<sup>46</sup> (Supplementary Figs. 27–31). We found 20 out of 21  $T_0$  lines contained edits at 13 or more sites with seven or more biallelic edits (Fig. 4c). Eleven  $T_0$  lines had edits at 15 sites, and one line (Line 2) had all 16 sites edited (Fig. 4c). The average editing and biallelic editing efficiencies for each target were 89.3% and 69.8%, respectively, indicating a high multiplexed editing frequency. These data suggest that multiplex system B can enable highly efficient large-scale biallelic genome editing in rice, with the ability to edit 16 sites and biallelically edit 14 sites simultaneously (Fig. 4d).

**Multiplexed transcriptional repression by compact STU dCas12a-SRDX systems.** Previously, we demonstrated transcriptional repression of single genes in plants with dCas12a-SRDX<sup>16,24</sup>. The STU systems represent the most compact expression systems for simultaneous and coordinated repression of multiple genes, which can potentially reduce the difficulties in cloning and transformation due to large constructs, as well as avoid gene silencing. Therefore, we chose two top performing STU systems, D and M (Fig. 3), for multiplexed transcriptional repression based on dLbCas12a-SRDX. First, we tested the systems in rice by simultaneously targeting four genes. For each system, two constructs were designed, and each construct contained four crRNAs targeting the promoters of four chosen genes (Supplementary Fig. 32a). Simultaneous transcriptional repression was observed in transformed rice protoplasts, especially when antisense strand was targeted with system M with the target genes repressed to ~20% of the control level (Supplementary Fig. 32b). We further tested the systems in *Arabidopsis* by targeting two tandemly arrayed genes, *At3g48090* and *At3g48080*, which encode two *EDS1* homologs (Supplementary Fig. 32c). Simultaneous transcriptional repression of both genes was observed in multiple  $T_1$  lines with either the D or M STU system (Supplementary Fig. 32d, e), and the repression effects were transmitted into the  $T_2$  generation (Supplementary Fig. 32f).

## Discussion

As a promising genome editing system, CRISPR-Cas12a is constantly under improvement in areas such as editing activities<sup>15,26</sup>, PAM requirements<sup>13,15</sup>, and multiplexing capabilities<sup>8,37</sup>. Only three Cas12a orthologs, LbCas12a, FnCas12a, and AsCas12a, have been previously applied for plant genome editing<sup>9,16–30</sup>. AsCas12a, which has low nuclease activities at lower temperatures<sup>24,43</sup>, could be improved by protein engineering<sup>15</sup> and directed evolution (e.g.,



**Fig. 4 Highly efficient large-scale biallelic genome editing in rice.** **a** Simultaneous editing of six sites by Mb2Cas12a for multiplexed engineering of quantitative traits in rice. Genotypes of four independent  $T_0$  lines are shown, indicating all the target sites are nearly edited biallelically. The PAM sites are highlighted in red and the target sequences are highlighted in blue. The TATA boxes are painted in red. The TAL effector binding sites in the promoters of the *SWEET* genes are indicated by solid black lines above the sequences. Exons of *GS3* and *GW2* are underlined. **b** Schematics of simultaneous editing by LbCas12a at 16 different target sites (T1–T16) across nine rice chromosomes (chr). **c** Genotypes of 21 tested  $T_0$  lines at 16 target sites, with wild type denoted as empty rectangles, monoallelic mutants denoted as half-filled rectangles, biallelic mutants denoted as fully filled rectangles, and chimeras denoted as dotted rectangles. **d** Genotypes of  $T_0$  Line 2 and Line 21 at 16 target sites, where 14 sites were biallelically edited in both lines.

Alt-R Cas12a (Cpf1) Ultra, IDT). Hence, it is important to identify more functional Cas12a orthologs, as they can directly or indirectly (e.g., through protein engineering) contribute to the expansion and improvement of the Cas12a toolbox. Toward this, four Cas12a orthologs, MbCas12a, Mb3Cas12a, HkCas12a, and ErCas12a, were recently demonstrated in mammalian cells<sup>14,39,41,44</sup>. In this study, we identified eight Cas12a orthologs that can mediate genome editing in rice. Six of them, ErCas12a, Lb5Cas12a, Bscas12a, Mb2Cas12a, Tscas12a, and MbCas12a, showed robust editing by generating biallelic mutants in rice. Mb2Cas12a can efficiently edit relaxed VTTV PAM sites, and its RVR variant can robustly edit TATV PAM sites in rice. Moreover, Mb2Cas12a-RVRR can edit

TTTTV, VTTTV, TATV, TYCV, CCCV, and CTCV PAM sites, greatly expanding the targeting scope. These functional Cas12a orthologs have also provided a rich Cas12a resource for future improvement through protein engineering.

Despite many previous efforts on developing multiplex Cas12a systems in plants<sup>9,25,53–57</sup>, a highly efficient multiplexed biallelic editing system remains elusive. For example, Cas12a mediated multiplexed editing was previously reported in rice at limited sites with biallelic editing frequencies up to 28%<sup>9</sup>, 33%<sup>57</sup>, 43%<sup>58</sup>, 67%<sup>55</sup>, respectively. By comparing 12 multiplexed Cas12a systems, here we successfully identified one multiplex system that enabled high-efficiency large-scale biallelic genome editing. This

system B uses a Pol II promoter to express a tandem HH-crRNA-HDV array. Using system B, we achieved simultaneous biallelic editing at 14 independent loci in the rice genome at a high frequency. We did not observe decreased editing efficiency when increasing the distance between the promoter and the crRNA (T1 is the closest and T16 is the farthest to the promoter) (Fig. 4c). We envision that its multiplexing capacity could be far beyond this, which requires future testing. Further, we applied two compact STU dCas12a-SRDX systems for multiplexed transcriptional repression in rice and *Arabidopsis*, providing an upgrade from singular transcriptional repression systems based on Cas12a<sup>16,24</sup>. Assembly of these useful multiplexed Cas12a systems is based on modular Golden Gate and Gateway vectors (Supplementary Table 4) and streamlined cloning strategies (Supplementary Fig. 33).

In conclusion, these systems have greatly expanded the targeting scope of Cas12a, and they collectively represent very powerful tools for crop breeding and engineering. A toolbox for using these Cas12a reagents has been established and is available to all researchers.

## Methods

**Plant material and growth condition.** Rice (*Oryza sativa*) plants used in this study were the Japonica cultivar *Nipponbare* and *Kitaake*. Seedlings of 14–16-day old grown on ½ MS medium<sup>59</sup> in dark at 28 °C were used for protoplast isolation<sup>60</sup>. Rice calli induced from mature rice embryos, were cultured on the N6D medium<sup>61</sup> under light at 32 °C and used for rice stable transformation. *Arabidopsis thaliana* (Col-0) plants grown at 22 °C were used for the gene repression study.

**Vector construction.** All vectors used in this study were constructed based on a three-way Gateway cloning system. T-DNA vectors used for single target editing were constructed using the previously described method<sup>60,62</sup>. To express Cas12a nuclease in attL1-attR5 entry vectors, three gBlocks were synthesized using rice codon optimized sequences for each Cas12a ortholog and assembled with the pYPQ230 (LbCas12a) backbone using the NEBuilder HiFi DNA Assembly Cloning Kit (New England BioLabs). To generate RVR and RVR variants, Mb2Cas12a-v1 and Mb2Cas12a-v2, mutations were introduced using the Q5<sup>®</sup> Site-Directed Mutagenesis Kit (New England BioLabs). To express crRNA in attL5-attL2 entry vectors, duplexed oligonucleotides were phosphorylated, annealed and ligated into pYPQ141-ZmUbi-RZ-Lb (for LbCas12a) and pYPQ141-ZmUbi-RZ-Fn (for all other Cas12a orthologs) at the *Esp3I* (*BsmBI*) site. These two types of entry vectors were further assembled with the destination vector pYPQ203<sup>63</sup> through LR reactions.

T-DNA vectors used for multiplexed genome editing and gene transcriptional regulation were constructed using the following method. The attL1-attR5 entry vectors included previously described pYPQ230 and pYPQ284 (Mb2Cas12a). To generate pYPQ230-STU to express Cas12a and its crRNAs as a single transcript unit (STU), the NOS terminator in pYPQ230 was replaced by a polyA signal at the *AatII* and *BspEI* sites. Similarly, to generate pYPQ233-STU, the NOS terminator in pYPQ233 (LbCas12a-SRDX) was replaced by a polyA signal. To express four crRNAs in attL5-attL2 entry vectors, the cloning method for each multiplexing system is as follows: (A) crRNA1 and 3 were cloned into pYPQ131C-RZ-Lb (Addgene #134347) while crRNA2 and 4 were cloned into pYPQ131D-RZ-Lb (Addgene #134348) at the *Esp3I* site<sup>33</sup>. All four crRNA expression cassettes were PCR amplified using primers flanked by *BsaI* recognition and cutting sites that are compatible with Golden Gate assembly. Four PCR products were assembled with recipient vector pYPQ144 (Addgene #69296)<sup>63</sup> using Golden Gate reactions. (B–D) To make Golden Gate recipient vector pYPQ144-pT, a polyT signal was inserted into pYPQ144 at the *SpeI* and *EcoRI* sites. To make Golden Gate recipient vector pYPQ144-ZmUbi-pT, pZmUbi was cut off from pYPQ141-ZmUbi-RZ-Lb with *AflIII* and *BamHI* and cloned into pYPQ144-pT. To make Golden Gate recipient vector pYPQ144C-pT, OsU6 promoter was PCR amplified from pYPQ131C (Addgene #69284)<sup>63</sup> and cloned into pYPQ144-pT at the *XbaI* and *BamHI* sites. The crRNA1–4 were then cloned into pYPQ131-STU-Lb/Fn, pYPQ132-STU-Lb/Fn, pYPQ133-STU-Lb/Fn, and pYPQ134-STU-Lb/Fn, respectively. These four vectors were assembled with recipient vectors pYPQ144-ZmUbi-pT, pYPQ144C-pT, and pYPQ144-pT using Golden Gate reactions for multiplexing systems B, C, and D, respectively. (E) To generate an empty vector for crRNA expression using the tRNA and ribozyme hybrid system, a gBlock containing OsU6 promoter, tRNA, crRNA direct repeat and cloning site, as well as the HDV ribozyme was synthesized and cloned into pYPQ131C at the *BamHI* and *NcoI* sites. This vector is named as pYPQ131C-tZ-Lb. The crRNA1–4 were then cloned into pYPQ131C-tZ-Lb, and all four crRNA expression cassettes without promoters were PCR amplified using primers flanked by *BsaI* recognition and cutting sites that are compatible with Golden Gate assembly. Four PCR products were assembled with recipient vector pYPQ144-pT using Golden Gate reactions. (F and G) To generate an empty vector for crRNA expression using the tRNA and ribozyme combined system,

a gBlock containing OsU6 promoter, tRNA, HH ribozyme, crRNA direct repeat and cloning site, as well as the HDV ribozyme was synthesized and cloned into pYPQ131C at the *BamHI* and *NcoI* sites. This vector is named as pYPQ131C-tDZ-Lb. The crRNA1–4 were then cloned into pYPQ131C-tDZ-Lb, and all four crRNA expression cassettes without promoters were PCR amplified using primers flanked by *BsaI* recognition and cutting sites that are compatible with Golden Gate assembly. Four PCR products were assembled with recipient vector pYPQ144C-pT and pYPQ144-pT using Golden Gate reactions for multiplexing systems F and G, respectively. (H) The CRISPR array driven by OsU6 promoter as well as a polyT signal at the end was synthesized and ligated into pYPQ144-pT at the *EcoRI* and *BamHI* sites. (M) The CRISPR array without promoter but with an extra crRNA direct repeat, and a polyT signal at the end was synthesized and ligated into pYPQ141-pT at the *EcoRI* and *BamHI* sites. (I and L) CRISPR array flanked by HH and HDV ribozymes as well as a polyT signal at the end was synthesized and ligated into pYPQ144-ZmUbi-pT and pYPQ141-pT at *EcoRI* and *BamHI* sites for multiplexing systems I and L, respectively. (J) CRISPR array flanked by HH and HDV ribozymes, driven by OsU6 promoter as well as a polyT signal at the end, was synthesized and ligated into pYPQ144-pT at the *EcoRI* and *BamHI* sites. The attL1-attR5 entry vectors used to edit six and 16 sites were assembled using the same method described previously for multiplexing system B, followed by a higher order assembly using restriction digestion and ligation. The attL1-attR5 entry vectors and attL5-attL2 entry vectors were further assembled with the destination vector pYPQ203 through LR reactions.

All vectors used for single target editing, as well as for multiplexed genome editing and gene transcriptional regulation using top performing multiplexing systems (B, D, I, L, and M), were deposited to Addgene. The Addgene plasmid numbers can be found in Supplementary Table 4. Information for all primers and gBlocks used in this study can be found in Supplementary Data 2. Information for all the T-DNA vectors used in this study can be found in Supplementary Data 3. A toolbox for multiplexed genome engineering using Cas12a has been established in this study and the detailed user manual is included in Supplementary Methods.

**Rice protoplast transformation.** Rice protoplasts were isolated and transformed according to a previously published protocol<sup>60</sup>. Briefly, 14–16-day old rice leaves grown in the dark were cut into 0.5–1.0 mm strips and incubated in the enzyme solution at 28 °C for 8 h without light. The digested cells were filtered by 75 µm cell strainer and washed by W5 buffer twice. 30 µg plasmid DNA was mixed with 200 µl protoplasts ( $2 \times 10^6 \text{ ml}^{-1}$ ) re-suspended in MMG buffer (400 µl protoplasts were used for the gene repression experiment). Equal amount of polyethylene glycol (PEG) transformation buffer was then added, and the entire mixture was incubated for 30 min at room temperature. The reactions were stopped by adding 900 µl W5 buffer. Protoplasts were collected by centrifugation and transferred into 12-well culture plates. Plates were incubated at 32 or 22 °C in dark for 2 days. The protoplasts were collected by centrifugation for DNA or RNA extraction. Protoplasts transformed with water were used as the wild type (WT) control for all genome editing experiments.

**Rice stable transformation.** Rice was transformed using the *Agrobacterium*-mediated method as described in the published protocols with slight modifications<sup>61,64,65</sup>. Briefly, *Agrobacterium tumefaciens* strain EHA105 harboring binary vectors was used to inoculate rice calli. Inoculated calli were co-cultured with *Agrobacterium* for 3 days, washed and moved to selection medium containing 50 mg l<sup>-1</sup> hygromycin. After 4 weeks, resistant calli were moved to regeneration medium I to induce shoot growth. Small shoots were further transferred to regeneration medium II to obtain full transgenic plants. DNA was extracted from young leaves of T<sub>0</sub> plants using the CTAB method for genotyping<sup>66</sup>.

**Mutation analysis by RFLP and SSCP.** To analyze mutation results using the restriction fragment length polymorphism (RFLP) method, the targeted genomic regions were amplified, and the PCR products were digested with restriction enzymes whose cutting sites are overlapping with the expected editing sites. Digested products were visualized on 2% TAE agarose gels. Mutation frequencies were quantified based on band intensity using Image Lab™ Software (Bio-Rad Laboratories) and ImageJ (<https://imagej.nih.gov/ij/>) for protoplast assays. For stable transgenic rice lines, any samples with undigested bands were considered edited, while any samples without detectable digested bands were considered biallelically edited in the RFLP analysis. To analyze mutation results in transgenic rice plants using the single-strand conformational polymorphism (SSCP) method<sup>67</sup>, amplicons were denatured for 5 min at 95 °C and immediately put on ice to minimize self-annealing. Denatured PCR amplicons were electrophoresed on 15 % non-denaturing polyacrylamide gels at 45 mA, 120–200 V. After 6 h, polyacrylamide gels were stained using argenteron.

**Sanger sequencing and deep sequencing.** PCR amplicons from stable transgenic rice were subjected to Sanger sequencing. DNA sequences were decoded using DSDecodeM<sup>68</sup>. PCR amplicons with sequencing barcodes generated from protoplast assay were first submitted for quality check, followed by sequencing using an Illumina HiSeqXten-PE150 system. Sequencing data were then analyzed by CRISPRMatch<sup>69</sup> for editing frequencies and profiles. For off-target analysis, amplicons from Line 10 (multiplexed genome editing by Mb2Cas12a targeting four sites) and Line 1

(multiplexed genome editing by Mb2Cas12a targeting six sites for quantitative traits) were pooled together, while amplicons from Line 11 (multiplexed genome editing by Mb2Cas12a targeting four sites) and Line 2 (multiplexed genome editing by Mb2Cas12a targeting six sites for quantitative traits) were pooled together. Potential off-target sites were sequenced using the Illumina HiSeq2500 system and analyzed using CRISPResso2<sup>70</sup>. To assess the editing results for multiplexed genome editing at 4 sites and 16 sites using system B, PCR amplicons were barcoded using the Hi-TOM<sup>46</sup> primers and pooled into two and four samples for Illumina HiSeq2500, respectively. Data were analyzed using the Hi-TOM online tool (<https://doi.org/www.hi-tom.net/hi-tom/>) with a 12.5% filter threshold.

**Arabidopsis stable transformation.** *Arabidopsis thaliana* wild type plants *Col-0* were transformed with *Agrobacterium tumefaciens* GV3101 using the flora dip method<sup>71</sup>. T<sub>1</sub> generation seeds were collected and sterilized with 50% bleach and 0.05% Tween. After 3 days of vernalization at 4 °C, seeds were plated on ½ MS medium supplemented with 15 mg l<sup>-1</sup> hygromycin. After a week, hygromycin resistant plants were transferred to ½ MS medium to recover for another week. Individual plants were treated as single transgenic lines. Plants transformed with a *GUS* gene and a hygromycin resistance gene were used as controls. Leaf tissue was collected from each line along with control plants for qRT-PCR analysis. Seeds collected from these lines were treated using the same method to obtain the T<sub>2</sub> generation plants. Three plants of each line were used for qRT-PCR analysis.

**Gene expression analysis using qRT-PCR.** Total RNA from rice and *Arabidopsis* leaf tissue or rice protoplast was extracted using TRIzol™ Reagent (Thermo Fisher Scientific) following the manufacturer's instructions with slight modifications. DNA was then removed with DNase I (New England Biolabs) and the complementary DNA (cDNA) was synthesized using the SuperScript™ III First-Strand Synthesis System (Thermo Fisher Scientific). Applied Biosystems™ SYBR™ Green PCR Master Mix (Thermo Fisher Scientific) was used for qRT-PCR. Three technical replicates were carried out for each sample using a CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad Laboratories). *AtEfl1a* was used as the reference gene for *Arabidopsis* and *OsTubulin* was used as the reference gene for rice. Protoplasts transformed with the T-DNA vector containing LbCas12a without crRNAs were used as controls for rice. Relative expression level of each gene was calculated using the 2<sup>-ΔΔC<sub>t</sub></sup> method.

**Genome-wide PAM analysis.** The rice cultivar Nipponbare genome sequence (*Oryza sativa nipponbare\_v7.0\_all.con*) was obtained from the Rice Genome Annotation Project ([rice.plantbiology.msu.edu/pub/data/Eukaryotic\\_Projects/o\\_sativa/annotation\\_dbs/pseudomolecules/](http://rice.plantbiology.msu.edu/pub/data/Eukaryotic_Projects/o_sativa/annotation_dbs/pseudomolecules/)). PAM sequence and its reverse complementary sequence were used to check the FASTA files using a regular expression search in Perl (version 5.26).

**Statistical analysis.** Pearson's  $\chi^2$  test was used to show whether there is a significant difference between the expected frequencies and the observed frequencies of edits in transgenic rice T<sub>1</sub> populations when  $\alpha = 0.05$ . Student's *t*-test was used for pairwise comparison. One asterisk ( $p < 0.05$ ) and two asterisks ( $p < 0.01$ ) indicate significant differences between two treatments. Tukey's Honest Significant Difference (HSD) test was used for multiple comparisons. Treatments with the same letter are not significantly different when  $\alpha = 0.05$ .

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

## Data availability

Next-generation sequencing (NGS) data have been deposited to the National Center for Biotechnology Information (NCBI) database under Sequence Read Archive (SRA) BioProject "PRJNA595844". The 35 Golden Gate and Gateway compatible vectors for the new Cas12a toolbox are available at Addgene (Supplementary Table 4). Source data are provided with this paper. Any other relevant data are available from the authors upon reasonable request. Source data are provided with this paper.

Received: 9 December 2020; Accepted: 8 March 2021;

Published online: 29 March 2021

## References

- Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).
- Zhang, Y., Malzahn, A. A., Sretenovic, S. & Qi, Y. The emerging and unutilized potential of CRISPR technology in plant science. *Nat. Plants* **5**, 778–794 (2019).
- Wu, W. Y., Lebbink, J. H. G., Kanaar, R., Geijsen, N. & van der Oost, J. Genome editing by natural and engineered CRISPR-associated nucleases. *Nat. Chem. Biol.* **14**, 642–651 (2018).
- Zhang, Y., Massel, K., Godwin, I. D. & Gao, C. Applications and potential of genome editing in crop improvement. *Genome Biol.* **19**, 210 (2018).
- Kim, D. et al. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol.* **34**, 863–868 (2016).
- Kleinstiver, B. P. et al. Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat. Biotechnol.* **34**, 869–874 (2016).
- Tang, X. et al. 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 (2018).
- Zetsche, B. et al. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat. Biotechnol.* **35**, 31–34 (2017).
- Wang, M., Mao, Y., Lu, Y., Tao, X. & Zhu, J. K. Multiplex gene editing in rice using the CRISPR-Cpf1 system. *Mol. Plant* **10**, 1011–1013 (2017).
- Yan, M. Y. et al. CRISPR-Cas12a-assisted recombineering in bacteria. *Appl. Environ. Microbiol.* **83**, e00947 (2017).
- Shen, W. et al. Establishment and application of a CRISPR-Cas12a assisted genome-editing system in *Zymomonas mobilis*. *Microb. Cell Factor.* **18**, 162 (2019).
- Li, L. et al. CRISPR-Cpf1-assisted multiplex genome editing and transcriptional repression in streptomyces. *Appl. Environ. Microbiol.* **84**, e00827–18 (2018).
- Gao, L. et al. Engineered Cpf1 variants with altered PAM specificities. *Nat. Biotechnol.* **35**, 789–792 (2017).
- Toth, E. et al. Mb- and FnCpf1 nucleases are active in mammalian cells: activities and PAM preferences of four wild-type Cpf1 nucleases and of their altered PAM specificity variants. *Nucleic Acids Res.* **46**, 10272–10285 (2018).
- Kleinstiver, B. P. et al. Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat. Biotechnol.* **37**, 276–282 (2019).
- Tang, X. et al. A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nat. Plants* **3**, 17018 (2017).
- Zhong, Z. et al. Plant genome editing using FnCpf1 and LbCpf1 nucleases at redefined and altered PAM sites. *Mol. Plant* **11**, 999–1002 (2018).
- Hu, X., Wang, C., Liu, Q., Fu, Y. & Wang, K. Targeted mutagenesis in rice using CRISPR-Cpf1 system. *J. Genet. Genomics* **44**, 71–73 (2017).
- Yin, X. et al. CRISPR-Cas9 and CRISPR-Cpf1 mediated targeting of a stomatal developmental gene EPFL9 in rice. *Plant Cell Rep.* **36**, 745–757 (2017).
- Begemann, M. B. et al. Precise insertion and guided editing of higher plant genomes using Cpf1 CRISPR nucleases. *Sci. Rep.* **7**, 11606 (2017).
- Xu, R. et al. Generation of targeted mutant rice using a CRISPR-Cpf1 system. *Plant Biotechnol. J.* **15**, 713–717 (2017).
- Li, S. et al. Precise gene replacement in rice by RNA transcript-templated homologous recombination. *Nat. Biotechnol.* **37**, 445–450 (2019).
- Lee, K. et al. Activities and specificities of CRISPR-Cas9 and Cas12a nucleases for targeted mutagenesis in maize. *Plant. Biotechnol. J.* **17**, 362–372 (2018).
- Malzahn, A. A. et al. Application of CRISPR-Cas12a temperature sensitivity for improved genome editing in rice, maize, and Arabidopsis. *BMC Biol.* **17**, 9 (2019).
- Bernabe-Orts, J. M. et al. Assessment of Cas12a-mediated gene editing efficiency in plants. *Plant Biotechnol. J.* **17**, 1971–1984 (2019).
- Schindele, P. & Puchta, H. Engineering CRISPR/LbCas12a for highly efficient, temperature-tolerant plant gene editing. *Plant Biotechnol. J.* **18**, 1118–1120 (2019).
- Kim, H. et al. CRISPR/Cpf1-mediated DNA-free plant genome editing. *Nat. Commun.* **8**, 14406 (2017).
- Li, B. et al. Robust CRISPR/Cpf1 (Cas12a)-mediated genome editing in allotetraploid cotton (*Gossypium hirsutum*). *Plant Biotechnol. J.* **17**, 1862–1864 (2019).
- Jia, H., Orbovic, V. & Wang, N. CRISPR-LbCas12a-mediated modification of citrus. *Plant Biotechnol. J.* **17**, 1928–1937 (2019).
- Endo, A., Masafumi, M., Kaya, H. & Toki, S. Efficient targeted mutagenesis of rice and tobacco genomes using Cpf1 from *Francisella novicida*. *Sci. Rep.* **6**, 38169 (2016).
- Li, S. et al. Synthesis-dependent repair of Cpf1-induced double strand DNA breaks enables targeted gene replacement in rice. *J. Exp. Bot.* **69**, 4715–4721 (2018).
- Wolter, F. & Puchta, H. In planta gene targeting can be enhanced by the use of CRISPR/Cas12a. *Plant J.* <https://doi.org/10.1111/tpj.14488> (2019).
- Li, S., Zhang, Y., Xia, L. & Qi, Y. CRISPR-Cas12a enables efficient biallelic gene targeting in rice. *Plant Biotechnol. J.* **18**, 1351 (2019).
- Li, X. et al. Base editing with a Cpf1-cytidine deaminase fusion. *Nat. Biotechnol.* **36**, 324–327 (2018).
- Tak, Y. E. et al. Inducible and multiplex gene regulation using CRISPR-Cpf1-based transcription factors. *Nat. Methods* **14**, 1163–1166 (2017).
- Breinig, M. et al. Multiplexed orthogonal genome editing and transcriptional activation by Cas12a. *Nat. Methods* **16**, 51–54 (2019).

37. Campa, C. C., Weisbach, N. R., Santinha, A. J., Incarnato, D. & Platt, R. J. Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. *Nat. Methods* **16**, 887–893 (2019).
38. Wang, L. et al. Improved CRISPR-Cas12a-assisted one-pot DNA editing method enables seamless DNA editing. *Biotechnol. Bioeng.* **116**, 1463–1474 (2019).
39. Zetsche, B. et al. A survey of genome editing activity for 16 Cas12a orthologs. *Keio J. Med.* **69**, 59–65 (2019).
40. Tu, M. et al. A ‘new lease of life’: FnCpf1 possesses DNA cleavage activity for genome editing in human cells. *Nucleic Acids Res.* **45**, 11295–11304 (2017).
41. Teng, F. et al. Enhanced mammalian genome editing by new Cas12a orthologs with optimized crRNA scaffolds. *Genome Biol.* **20**, 15 (2019).
42. Li, S. et al. Expanding the scope of CRISPR/Cpf1-mediated genome editing in rice. *Mol. Plant* **11**, 995–998 (2018).
43. Moreno-Mateos, M. A. et al. CRISPR-Cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing. *Nat. Commun.* **8**, 2024 (2017).
44. Wierson, W. A. et al. Expanding the CRISPR Toolbox with ErCas12a in zebrafish and human cells. *CRISPR J.* **2**, 417–433 (2019).
45. Xie, K., Minkenberg, B. & Yang, Y. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl Acad. Sci. USA* **112**, 3570–3575 (2015).
46. Liu, Q. et al. Hi-TOM: a platform for high-throughput tracking of mutations induced by CRISPR/Cas systems. *Sci. China Life Sci.* **62**, 1–7 (2019).
47. Mao, H. et al. Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. *Proc. Natl Acad. Sci. USA* **107**, 19579–19584 (2010).
48. Song, X. J., Huang, W., Shi, M., Zhu, M. Z. & Lin, H. X. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* **39**, 623–630 (2007).
49. Oliva, R. et al. Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat. Biotechnol.* **37**, 1344–1350 (2019).
50. Eom, J. S. et al. Diagnostic kit for rice blight resistance. *Nat. Biotechnol.* **37**, 1372–1379 (2019).
51. Li, T., Liu, B., Spalding, M. H., Weeks, D. P. & Yang, B. High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat. Biotechnol.* **30**, 390–392 (2012).
52. Zhou, J. et al. Multiplex QTL editing of grain-related genes improves yield in elite rice varieties. *Plant Cell Rep.* **38**, 475–485 (2018).
53. Ding, D., Chen, K., Chen, Y., Li, H. & Xie, K. Engineering introns to express RNA guides for Cas9- and Cpf1-mediated multiplex genome editing. *Mol. Plant* **11**, 542–552 (2018).
54. Xu, R. et al. Enhanced genome editing in rice using single transcript unit CRISPR-LbCpf1 systems. *Plant Biotechnol. J.* **17**, 553–555 (2019).
55. Wang, M. et al. Multiplex gene editing in rice with simplified CRISPR-Cpf1 and CRISPR-Cas9 systems. *J. Integr. Plant Biol.* **60**, 626–631 (2018).
56. Zhong, Z. et al. Intron-based single transcript unit CRISPR systems for plant genome editing. *Rice* **13**, 8 (2020).
57. Tang, X. et al. Single transcript unit CRISPR 2.0 systems for robust Cas9 and Cas12a mediated plant genome editing. *Plant Biotechnol. J.* **17**, 1431–1445 (2019).
58. Hu, X., Meng, X., Li, J., Wang, K. & Yu, H. Improving the efficiency of the CRISPR-Cas12a system with tRNA-crRNA arrays. *Crop J.* **8**, 403–407 (2019).
59. Murashige, T. & Skoog, F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant Physiol.* **15**, 473–497 (1962).
60. Tang, X. et al. A single transcript CRISPR-Cas9 system for efficient genome editing in plants. *Mol. Plant* **9**, 1088–1091 (2016).
61. Toki, S. Rapid and efficient Agrobacterium-mediated transformation in rice. *Plant Mol. Biol. Rep.* **15**, 16–21 (1997).
62. Zhang, Y., Zhang, Y. & Qi, Y. Plant gene knockout and knockdown by CRISPR-Cpf1 (Cas12a) systems. *Methods Mol. Biol.* **1917**, 245–256 (2019).
63. Lowder, L. G. et al. A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. *Plant Physiol.* **169**, 971–985 (2015).
64. Armstrong, C. L. & Green, C. E. Establishment and maintenance of friable, embryogenic maize callus and the involvement of L-proline. *Planta* **164**, 207–214 (1985).
65. Hiei, Y., Ohta, S., Komari, T. & Kumashiro, T. Efficient transformation of rice (*Oryza sativa* L.) mediated by Agrobacterium and sequence analysis of the boundaries of the T-DNA. *Plant J.* **6**, 271–282 (1994).
66. Murray, M. G. & Thompson, W. F. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* **8**, 4321–4325 (1980).
67. Zheng, X. et al. Effective screen of CRISPR/Cas9-induced mutants in rice by single-strand conformation polymorphism. *Plant Cell Rep.* **35**, 1545–1554 (2016).
68. 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).
69. 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).
70. Clement, K. et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **37**, 224–226 (2019).
71. Clough, S. J. & Bent, A. F. Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).

## Acknowledgements

We thank Dr. Brian Jaffaldano and Simon Sretenovic for the preparation of the manuscript. This work was supported by Syngenta, Foundation for Food and Agriculture Research grant (award no. 593603), the National Science Foundation Plant Genome Research Program grant (award no. IOS-1758745), Biotechnology Risk Assessment Grant Program competitive grant (award no. 2018-33522-28789) from the U.S. Department of Agriculture, and University of Maryland startup funds to Y.Q. It was also supported by the National Transgenic Major Project (award no. 2018ZX08020-003), the National Natural Science Foundation of China (award no. 31771486 and 31960423), the Fundamental Research Funds for the Central Universities (award no. ZYGX2019J127) and the Science Strength Promotion Program of UESTC to Yong Z. and X.Z. Y.C. was supported by a scholarship from China Scholarship Council. The content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of the Foundation for Food and Agriculture Research.

## Author contributions

Y.Q., Yong Z., and Q.Q. conceived and supervised the research. Y.Q., Yong Z., and Yingxiao Z. designed the experiments. Yingxiao Z. made all the vectors. Yingxiao Z., Q.R., S.L., and X.T. did rice protoplast transformation and analysis. Yingxiao Z., Q.R., X.T., S.L., J.Z., J.W., X.Z., D.Y., L.T., M.Y., L.H., H.Y., Y.Z., Q.F., X. Z., Y.L., and B.M. performed rice stable transformation and genotyping analysis. A.M. and L.F. did the transcriptional repression experiment in *Arabidopsis*. P.C. did the transcriptional repression experiment in rice. J.S. and S.M. did the genome-wide PAM analysis in rice, maize, and wheat. Yingxiao Z. and Y.C. examined *SWEET* gene expression in edited rice lines. Y.Q., Yingxiao Z., and Yong Z. wrote the paper with input from other authors. All authors read and approved the final manuscript.

## Competing interests

The authors declare the following competing interests: Yingxiao Z. and Y.Q. are inventors on a U.S. Patent Application (No. 17/090,766) that has been filed on plant genome editing with Cas12a orthologs and multiplexed editing systems reported in this study. All other authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-22330-w>.

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

**Peer review information** *Nature Communications* thanks Nian Wang for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021