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

Yingxiao Zhang1,9, Qiurong Ren2,9, Xu Tang2,9, Shishi Liu2, Aimee A. Malzahn1, Jianping Zhou2, Jiaheng Wang2, Desuo Yin1,3, Changtian Pan1, Mingzhu Yuan2, Lan Huang2, Han Yang2, Yuxin Zhao2, Qing Fang2, Xuelian Zheng2, Li Tian2, Yanhao Cheng1,4, Ysa Le1, Bailey McCoy1, Lidiya Franklin1, Jeremy D. Selengut5, Stephen M. Mount6, Qiudeng Que7, Yong Zhang2 & Yiping Qi1,8

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.
CRISPR-Cas12a (formerly Cpf1) is a class II type V endonuclease that prefers a thymine-rich protospacer adjacent motif (PAM)\(^1\), and is the second most commonly used CRISPR system for genome editing\(^2\)–\(^4\). Cas12a has shown higher targeting specificity than Cas9 in mammalian cells and plants\(^5\)–\(^7\). 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\(^8\)–\(^9\). Cas12a has been used to generate targeted mutations as well as achieve transcriptional regulation in microorganisms\(^10\)–\(^12\). Cas12a can also efficiently generate edits in some important industrial Streptomyces strains that cannot be edited using SpCas9 due to toxicity\(^12\). In mammalian systems, Cas12a has been widely used and engineered for genetic manipulation\(^13\)–\(^15\). Three Cas12a orthologs, LbCas12a, AsCas12a, and FnCas12a, have been used in plants. LbCas12a is most popular due to its high editing activity in rice\(^16\)–\(^22\), maize\(^23\)–\(^24\), Arabidopsis\(^25\)–\(^26\), tomato\(^25\), Nicotiana benthamiana\(^15\), lettuce\(^27\), cotton\(^28\), and citrus\(^29\). FnCas12a also mediates efficient genome editing in plants\(^9\)–\(^17\),\(^20\),\(^23\)–\(^30\). 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\(^20\),\(^22\),\(^23\)–\(^31\). Furthermore, Cas12a has been repurposed for base editing\(^32\),\(^34\) and transcriptional regulation\(^15\)–\(^17\),\(^29\)–\(^35\),\(^37\)–\(^39\).

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 TTTV (V = A, C, and G) PAM requirement of Cas12a limits its targeting scope. Many attempts have been made to engineer Cas12a orthologs and engineered their variants in rice for efficient genome editing\(^4\),\(^10\)–\(^12\). Cas12a orthologs can induce medium to high-frequency genome editing in rice\(^9\),\(^17\),\(^20\),\(^23\)–\(^30\). 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\(^20\),\(^22\),\(^23\)–\(^31\). Furthermore, Cas12a has been repurposed for base editing\(^15\)–\(^17\),\(^29\)–\(^35\),\(^37\)–\(^39\).

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, MiCas12a, BoCas12a, MbCas12a, and CMaCas12a\(^1\)–\(^39\). In addition, we included ErCas12a\(^14\), which is also known as MAD7\(^{30}\). 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\(^36\) (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 efficiencies to LbCas12a. Targeted mutations were predominately deletions distal from PAMs, centered around the predicted cleavage site\(^1\) (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 efficiencies (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-freqency genome editing in rice except for MiCas12a and BoCas12a. To assess the inheritance of targeted mutations, we examined Mb2Cas12a T\(_2\) 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\(^17\), 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\(^17\). 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 T0 lines. f Editing analysis of transgenic rice T1 lines. χ² test results indicate that T1 populations from each tested T0 line segregated at a 1:2:1 ratio when α = 0.05. Data in “b–d” are presented as mean values ± SEM. n = 3 biologically independent samples. Source data are provided as a Source Data file.
engineered to target TATV PAMs in mammalian cells. We showed that LbCas12a-RVR could only target TATG PAMs while FnCas12a-RVR had no activity at TATV PAMs in rice. 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 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 RVR variant of Mb2Cas12a has a broad target range. To further broaden the target range of Mb2Cas12a, we engineered the RVR 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 TTTV 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 T0 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 α = 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, 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 ± 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, TTV, TATV, TTYC, CCCV, and CTCV PAM sites (Y = T and C) (Fig. 2e). The variant however showed no or low 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 temperatures24,43. 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)15, and found Mb2Cas12a-v2 had improved activity at the OsEPFL9-TTTG PAM site at 22 °C (Supplementary Fig. 12). These data suggest Mb2Cas12a-RVRR 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, TTV, TATV, TTYC, CCCV, and CTCV PAM sites (Y = T and C) (Fig. 2e). The variant however showed no or low 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.

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 promoter activity43. Given tRNA processing may leave a few extra nucleotides at the 5' end45. 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 exploited3. 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 T0 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 T0 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 OsUs6 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 humans56 and plants58; 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, 33, and 36 T0 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 T0 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-
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 multiplexed 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).
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 T₀ 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) and OsGW2 (GRAIN WIDTH and WEIGHT2) 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, creating bacterial leaf blight. 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, AvrXa7, and PthXo3 (Fig. 4a).

High biallelic editing efficiencies at all six target sites in eight out of 10 individual T₀ 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) and Cas9. Simultaneous editing of OsGS3 and OsGW2 was also achieved in these lines (Fig. 4a), which is expected to contribute to increased yield. Off-target analysis of two multiplex edited T₀ 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 T₀ lines with RFLP assays and the Hi-TOM sequencing platform. We found 20 out of 21 T₀ lines contained edits at 13 or more sites with seven or more biallelic edits (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. 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 T₀ lines with either the D or M STU system (Supplementary Fig. 32d, e), and the repression effects were transmitted into the T₂ generation (Supplementary Fig. 32f).

Discussion
As a promising genome editing system, CRISPR-Cas12a is constantly under improvement in areas such as editing activities, PAM requirements, and multiplexing capabilities. Only three Cas12a orthologs, LbCas12a, Fncas12a, and AsCas12a, have been previously applied for plant genome editing. AsCas12a, which has low nuclease activities at lower temperatures, could be improved by protein engineering and directed evolution (e.g.,...
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 cells14,39,41,44. 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 TTTV, VTTV, 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 plants9,25,53–57, 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%57, 33%57, 43%58, 67%55, respectively. By comparing 12 multiplexed Cas12a systems, here we successfully identified one multiplex system that enabled high-efficiency large-scale biallelic genome editing.

**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<sub>0</sub> 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<sub>0</sub> lines at 16 target sites, with wild type denoted as empty rectangles, monallelic mutants denoted as half-filled rectangles, biallelic mutants denoted as fully filled rectangles, and chimeras denoted as dotted rectangles. d Genotypes of T<sub>0</sub> Line 2 and Line 21 at 16 target sites, where 14 sites were biallelically edited in both lines.

TTTV, VTTT, 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.
system B uses a Pol II promoter to express a tandem HH-cRNA-HDV array. Using system B, we achieved simultaneous biallelic transcriptional repression in rice and Arabidopsis, providing an upgrade from singular transcriptional repression systems based on Cas12a. 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 14-day-old seedlings of 14–20 days after flowering. Seedsings of 14–20 days after flowering grown on ½ MS medium in dark at 28 °C were used for protoplast isolation. Rice calli induced from mature rice embryos, were cultured on the N6 medium 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. To express Cas12a nuclease in attL5-attR2 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-v and Mb2Cas12a-v2, modifications were introduced using the QuikChange Site-Directed Mutagenesis Kit (New England Biolabs). To express crRNA in attL5-attR2 entry vectors, diode 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 EspI (BsmBI) cloning site. These two types of entry vectors were further assembled with the destination vector pYPQ293 through LR reactions.

T-DNA vectors used for multiplexed genome editing and gene transcriptional regulation using top performing multiplexing systems (B, D, L, I, M, and G) 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. For all multiplexed systems, vectors used in this study were 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. 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 5 W buffer twice. 30 µg plasmid DNA was mixed with 200 µl protoplasts (2 × 10^6 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 at 4 °C for 30 min at room temperature. The reaction was then annealed 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 transformation was achieved using the Agrobacterium-mediated method as described in the published protocols with slight modifications. 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 hygromycin. After 4–6 weeks of resistant calli were transferred to regeneration medium I to induce shoot growth. Slow shoots were further transferred to regeneration medium II to obtain full transgenic plants. DNA was extracted from young leaves of T0 plants using the CTAB method for genotyping.

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 Image (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 biologically edited in the RFLP analysis. To analyze mutation results in transgenic rice plants using the single-strand conformational polymorphism (SSCP) method, amplicons were denatured for 5 min at 95 °C and immediately put on ice to minimize self-annexing. 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 argentation.

Sanger sequencing and deep sequencing. PCR amplicons from stable transgenic rice were subjected to Sanger sequencing. DNA sequences were decoded using Dn赛后Decode®. PCR amplicons with sequencing barcodes generated from protoplast assay were first submitted for quality check, followed by sequencing using an Illumina HiSeq2500 system. Four PCR products were then assembled usingSelectXMatch® for editing frequencies and profiles. For off-target analysis, amplicons from Line 10 (multiplexed genome editing by Mb2Cas12a targeting four sites) and Line 1...
DNA was then removed with DNase I (New England BioLabs) and the com-
using the 2
were used as controls for rice. Relative expression level of each gene was calculated

Statistical analysis

Gene expression analysis using qRT-PCR

Embryo rescue transgenic line plants of C. annuum were generated by Agrobacterium tumefaciens GV3101 using the CRISPR-Cas12a system. A total of 50 transgenic plants were generated. Three transgenic lines were used for qRT-PCR analysis. Leaf tissue was collected from each line along with controls plants for qRT-PCR analysis. Leaf tissue was collected from each line along with controls plants for qRT-PCR analysis.

Tang, X. et al. A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cas9 (Cas12a) nucleases in rice. 

Arabidopsis stable transformation. Arabidopsis thaliana wild type plants Col-0 were transformed with Agrobacterium tumefaciens GV3101 using the floral dip method17. T1 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 1/2 MS medium supplemented with 15 mg/l hygromycin. After a week, hygromycin resistant plants were transferred to 1/2 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. Leaf tissue was collected from each line along with control plants for qRT-PCR analysis.

Tang, X. et al. A CRISPR-Cas12a system for efficient genome editing and transcriptional repression in plants. 

Data availability

Individual plants were treated as single transgenic lines. Plants transformed with a

References

1. Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 163, 759–771 (2015).

3. Wu, W. Y., Lebbink, J. H. G., Kanaar, R., Geijtenbeek, N. & van der Oost, J. Genome editing by natural and engineered CRISPR-associated nucleases. Nat. Chem. Biol. 14, 642–651 (2018).

2. Zhang, Y., Massel, K., Godwin, I. D. & Gao, C. Applications and potential of genome editing in crop improvement. Genome Biol. 19, 210 (2018).

4. Kim, D. et al. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. Nat. Biotechnol. 34, 863–868 (2016).

5. Kleinstiver, B. P. et al. Genome-wide specificities of CRISPR-Cas9 nucleases in human cells. Nat. Biotechnol. 34, 869–874 (2016).

6. Tang, X. et al. A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cas9 (Cas12a) nucleases in rice. Genome Biol. 19, 84 (2018).

7. Zetsche, B. et al. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. Nat. Biotechnol. 33, 175–177 (2015).

8. 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).

9. Yan, M. Y. et al. CRISPR-Cas12a-assisted recombineering in bacteria. Appl. Environ. Microbiol. 83, e00947 (2017).

10. Shem, W. et al. Establishment and application of a CRISPR-Cas12a assisted genome-editing system in Zymomonas mobilis. Microb. Cell Factor 18, 162 (2019).

11. Li, L. et al. CRISPR-Cpf1-assisted multiplex genome editing and transcriptional repression in streptomycetes. Appl. Environ. Microbiol. 84, e00827–18 (2018).

12. Gao, L. et al. Engineered Cpf1 variants with altered PAM specificities. Nat. Biotechnol. 35, 789–792 (2017).

13. 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).

14. 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, 276–286 (2019).

15. Tang, X. et al. A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. Nat. Plants 3, 17018 (2017).

16. Zhong, Z. et al. Plant genome editing using FnCpf1 and LbCpf1 nucleases at redefined and altered PAM sites. Mol. Plant 11, 999–1002 (2018).

17. Hu, X., Wang, C., Liu, Q., Fu, Y. & Wang, K. Targeted mutagenesis in rice using a new CRISPR-Cpf1 system. J. Genet. Genomics 44, 71–73 (2017).

18. Yin, X. et al. CRISPR-Cas9 and CRISPR-Cpf1 mediated targeting of a somatic developmental gene EPFL9 in rice. Plant Cell Rep. 36, 745–757 (2017).

19. Begemann, M. B. et al. Precise insertion and guided editing of higher plant genomes using Cpf1 CRISPR nucleases. Sci. Rep. 7, 11606 (2017).

20. Xu, R. et al. Generation of targeted mutant rice using a CRISPR-Cpf1 system. Plant Biotechnol. J. 15, 713–717 (2017).

21. Li, S. et al. Precise genome replacement in rice by RNA transcript-templated homologous recombination. Nat. Biotechnol. 37, 445–450 (2019).

22. 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).

23. 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).

24. Bernabe-Orts, J. M. et al. Assessment of Cas12a-mediated gene editing efficiency in plants. Plant Biotechnol. J. 17, 1971–1984 (2019).

25. Schindelé, P. & Puchta, H. Engineering CRISPR/LbCas12a for highly efficient, temperature-tolerant plant gene editing. Plant Biotechnol. J. 18, 1118–1120 (2019).

26. Kim, H. et al. CRISPR/Cpf1-mediated DNA-free plant genome editing. Nat. Commun. 8, 14406 (2017).

27. Li, B. et al. Robust CRISPR/Cpf1 (Cas12a)-mediated genome editing in allotetraploid cotton (Gossypium hirsutum). Plant Biotechnol. J. 17, 1862–1864 (2019).

28. Jia, H., Orbovic, V. & Wang, N. CRISPR/LbCas12a-mediated modification of citrus. Plant Biotechnol. J. 17, 1928–1937 (2019).

29. 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).

30. Li, S. et al. Synthesis-dependent repair of Cpf1-induced double strand DNA breaks enables targeted genome replacement in rice. J. Exp. Bot. 69, 4715–4721 (2018).

31. Wolter, F. & Puchta, H. In plants gene targeting can be enhanced by the use of CRISPR/Cas12a. Plant J. https://doi.org/10.1111/tpj.14488 (2019).

32. Li, S., Zhang, Y., Xia, L. & Qi, Y. CRISPR-Cas12a enables efficient biallelic gene targeting in rice. Plant Biotechnol. J. 18, 1351 (2019).

33. Li, X. et al. Base editing with a Cpf1-cytidine deaminase fusion. Nat. Biotechnol. 36, 324–327 (2018).

34. Tak, Y. E. et al. Inducible and multiplex gene regulation using CRISPR-Cpf1-based transcription factors. Nat. Methods 14, 1163–1166 (2017).

35. Breising, 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., Santinina, 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. Nat. Med. 25, 897–901 (2019).
40. Tu, M. et al. A ‘new lease of life’: FncCpf1 possesses DNA cleavage activity for genome editing in human cells. Nucleic Acids Res. 45, 11295–11304 (2017).
41. Deng, 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 systems. Sci. China Life Sci. 62, 1–7 (2019).
47. Mao, H. et al. Linking differential domain functions of the GS3 protein to the boundaries of the T-DNA. Mol. Plant 12, 1372–1379 (2019).
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. 49, 623–630 (2017).
49. Obwa, 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, 1572–1579 (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. Zhang, 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-cRNA 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 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 knockdown 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 transposon regulatory. 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. Hiie, Y., Ohtia, 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. DSEdecide: 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).