Check for updates

JIPB Journal of Integrative Plant Biology

Multiplex gene editing in rice with simplified CRISPR-Cpf1 and CRISPR-Cas9 systems[FA]

Letter to the Editor

High-Impact Article

Summary We developed simplified single transcriptional unit (SSTU) CRISPR systems for multiplex gene editing in rice using FnCpf1, LbCpf1 or Cas9, in which the nuclease and its crRNA array are co-expressed from a single Pol II promoter, without any additional processing machinery. Our SSTU systems are easy to construct and effective in mediating multiplex genome editing.

The class 2 clustered regularly interspaced short palindromic repeat (CRISPR) systems have been widely used for simultaneous modification of multiple loci in plants. Traditionally, the type II CRISPR-Cas9 or type V CRISPR-Cpf1 (also known as Cas12a) system is a two-component transcriptional unit (TCTU) in which the Cas9 or Cpf1 protein is expressed from an RNA polymerase (pol) II promoter, whereas the single guide RNA (sgRNA) is typically expressed from a Pol III promoter, such as U6 or U3 promoter. For multiplex gene editing, a single sgRNA targeting multiple genes can be used in some cases (Yu et al. 2018) but most often multiple sgRNAs are needed, which can be driven by separate promoters (Ma et al. 2015; Zhang et al. 2016). Alternatively, the multiple sgRNAs can be expressed as a single transcript, which generates functional individual sgRNAs after processing by exogenous ribozymes (Gao and Zhao 2014), Csy4 (Čermák et al. 2017) or the plant endogenous transfer RNA (tRNA) processing system (Xie et al. 2015).

To simplify multiplex gene editing in plants, single transcriptional unit (STU) CRISPR-Cas9 systems were tested previously, in which Cas9, sgRNAs and ribozyme or tRNA elements were co-expressed from a single Pol II promoter (Tang et al. 2016; Ding et al. 2018). The CRISPR-Cpf1 system with a simple short direct repeat (DR)-guide array can also simplify multiplex gene editing, owing to the CRISPR RNA (crRNA) processing activity of the Cpf1 protein itself (Wang et al. 2017). However, compared to Cas9, the Cpf1 system has not been widely used for either single or multiple gene editing, possibly due to its low efficiency in the conventional pol III-derived TCTU system.

Here, we developed a simplified single transcriptional unit (SSTU) CRISPR system for multiplex gene editing in rice using FnCpf1, LbCpf1 or Cas9, in which the nuclease and its crRNA array are co-expressed from a single Pol II promoter, without any additional processing machinery. Our SSTU systems are simple and efficient in mediating multiplex genome editing for majority of the tested targets.

Previously, we demonstrated high efficiency multiplex editing in four genes by FnCpf1 and LbCpf1, using a simple short DR-guide array driven by the *Os*U6 promoter (Wang et al. 2017). With the same strategy, we tried to simultaneously edit eight genes in the rice late embryogenesis abundant (LEA) family (sub-family LEA_1 and LEA_2) using FnCpf1 (TCTU-1 in Figure S1). FnCpf1 recognizes TTN PAM that is more flexible for guide RNA designing in rice than the TTTN PAM required for LbCpf1. Unexpectedly, we observed that the editing efficiency was very low, only low mutation rates were observed at two of the target loci (Table S1).

It was reported that the efficiency of a crRNA expressed from a Pol III transcript decreases as the length of the crRNA array increases in mammalian cells (Zhong et al. 2017). We speculated that the *Os*U6 promoter may not be strong enough to efficiently drive a long crRNA array. Thus, we split the crRNA array into two parts, each consisting of four DR-guide units and driven by an individual *Os*U6 promoter (TCTU-2 in Figure S1). We found that the mutation rates at loci 1,258, 6,262 and 4,998 increased significantly compared to those of TCTU-1, but the rates remained very low at the other five targets (Table S1).

Given our previous study showing that four mature DR-guides driven by the *Os*U6 promoter were highly efficient in multiplex gene editing, we speculated that the low efficiency at the five targets above might be due

© 2018 Institute of Botany, Chinese Academy of Sciences

to their PAM-guide sequences. Taken together, these results suggest that modifications of both the crRNA expression cassette and PAM-guide may be necessary for further improvement in multiplex gene editing by Cpf1.

To optimize crRNA expression, different configurations of crRNA expression cassettes, driven by a pol III or pol II promoter, were tested for the *OsPDS-2* target site, which showed a very low mutation efficiency (1.8%) in our previous study (Wang et al. 2017). We determined that the use of the maize *ubiquitin* (*Zm*Ubi) promoter plus hammerhead and HDV ribozyme resulted in significantly higher mutation rate than the other cassettes (Figure S2). Our result is consistent with the study of Tang et al. (2017), and therefore, this configuration was further adopted for multiplex gene editing.

FnCpf1 was first reported to recognize the TTN PAM (Zetsche et al. 2015). However, a later study demonstrated that to achieve high cleavage activity in human cells, FnCpf1 requires a PAM defined as 5'-YTV-3' (Y represents C or T, and V represents A, C or G), and the spacer sequence with a length of 21 nt, rather than the commonly used 23-25 nt, could lead to maximal editing (Tu et al. 2017). Another study further showed that FnCpf1 preferred the TTTV motif to cleave target DNA when expressed in *Saccharomyces cerevisiae* (Świat et al. 2017). Based on these reports, we modified some of the guides and cloned the array into the TCTU expression cassette, together with hammerhead and HDV ribozyme (Figure 1A; Table S2).

After analyses of the T0 transgenic plants, we confirmed high editing efficiencies (>50%) at all four targets (4,353, 5,071, 2,387 and 2,191) with modified TTTV PAM-guides, and comparable or higher efficiencies at targets 1,258, 6,262 and 4,998 with unmodified PAM-guides. However, no mutation was detected in target 0204, for which the same TTN PAM was used and the guide was shortened from 25 nt to 22 nt (Figure 1A; Table S2). These results indicated that a TTTV PAM is beneficial for the efficiency of editing by FnCpf1. More recently, another group reported similar results in rice protoplasts (Zhong et al. 2018).

Since both FnCpf1 and its crRNA array were driven by a *Zm*Ubi promoter, we tried to combine them into a single transcript by inserting the array sequence into the 3' UTR of the FnCpf1 expression cassette, to form a SSTU system. We established that this simplified system gave comparable mutagenesis efficiencies with those of the TCTU system at targets 4,353, 5,071, 0,204 and 2,387, although the efficiencies were lower at the other targets (Figure 1A; Table S2). Elimination of one *Zm*Ubi promoter from TUTU was not only significantly reducing the vector size, but also avoiding potential risks from the use of duplicate promoters in a single T-DNA, such as gene silencing and recombination of the construct. We also noticed that in Figure S2, although DR sequences were set at both the 5' and 3' of the guide, in which the functional DR-guide unit theoretically can be fully released by the Cpf1 protein, none of the pol II promoter cassettes gave high editing efficiency when additional processing machinery was not included. However, our results suggest that this kind of RNA processing machinery is not necessary in the SSTU construction.

We further compared the TCTU and SSTU systems, using LbCpf1 to target another *Os*LEA sub-family, LEA_Dehydrins, which include eight members. We designed two target sites adjacent to locus 4,528 for deleting the whole gene, since no TTTV PAM exists in its CDS region. In this experiment, we chose the *Cestrum yellow leaf curling virus* promoter (CmYLCV) for crRNA array expression in the TCTU system to prevent the use of duplicate promoters (Figure 1B). The CmYLCV promoter drives comparable or higher levels of expression than the 35S or *Zm*Ubi promoter, and it has been applied successfully for multiplex gene editing in tomato cells, together with $tRAN^{Gly}$ elements (Čermák et al. 2017).

As shown in Figure 1B and detailed in Table S3, this SSTU system was comparable with TCTU at causing mutagenesis at eight of the nine tested target sites, although the efficiency was much lower at target 2,675. A series of mutants was produced from the TCTU and SSTU systems using both FnCpf1 and LbCpf1, although no octuple mutants were obtained, due to low efficiencies at some of the target sites (Tables S4–S7). Improved computational algorithms are needed for better prediction of Cpf1 guide RNA activities.

In previous reports of STU systems for Cas9 in plants, special processing devices, such as the hammerhead ribozyme (Tang et al. 2016) or tRNA element (Ding et al. 2018), was used to release individual sgRNAs. Recently, it was shown that an RNA transcript, consisting of a sgRNA adjoining a GFP protein coding region in the *Tobacco mosaic virus*, caused target indels



Figure 1. Multiplex gene editing in rice using TCTU and SSTU systems containing CRISPR-Cpf1 or CRISPR-Cas9
(A) T-DNA constructs of TCTU and SSTU systems for multiplex gene editing with FnCpf1 and the resulting mutagenesis efficiency at rice LEA_1 and LEA_2 family genes. **(B)** T-DNA constructs of TCTU and SSTU systems for multiplex gene editing with LbCpf1 and the resulting mutagenesis efficiency at rice LEA_Dehydrin family genes.

and viral-based GFP overexpression, simultaneously (Cody et al. 2017). Inspired by this finding, we speculated that the Cas9 protein and its sgRNAs may be assembled into a SSTU system, without any additional processing machinery. More recently, it has also been shown that the SpCas9 protein and intrinsic plant RNase, can produce functional sgRNA from a fused SpCas9-gRNA primary transcript, such that the translated SpCas9 protein can be bound to this transcript, and the extra RNA sequences are trimmed by RNase III and/or RNase T1 (Mikami et al. 2017).

In our design, three sgRNAs targeting loci 230, 250 and 240 in the rice CYP81A family were tandemly ligated into the 3' UTR of the Cas9 expression cassette, separated only by a 6 bp linker (the linker can contain a restriction site for cloning). A single sgRNA expression cassette for targeting locus 260 in the same family was designed as a control to evaluate the activity of the Cas9 protein in the SSTU system (Figure 1C). To compare the efficiency of the sgRNA array, a traditional multiplex gene editing system was also constructed, in which these three sgRNAs were driven by separate pol III promoters (TCTU-3). The SSTU system yielded good mutagenesis efficiency at the tested targets, except for target 250 (Figure 1C; Tables S8–S10).

The synchronized expression of Cas9 and sgRNAs in one cassette in the SSTU is presumably advantageous for efficient gene editing. The lower mutation rate observed for the middle sgRNA unit, targeting locus 250, may be a consequence of imprecise RNase trimming of the sgRNA array. Using longer linkers may help resolve this problem. Most of the mutations induced by Cas9 were homozygous and bi-alleles, in contrast to higher proportion of heterozygous and chimeric mutants induced by Cpf1 (Figures S3–S12; Tables S1–S10).

The strategies for multiplex gene editing in plants have evolved from the traditional TCTU systems to STU systems. In this study, we further simplified the STU architecture to a SSTU system for both Cpf1 and Cas9, without any additional processing machinery. Overall, our SSTU systems are simple to construct and effective at mediating multiplex genome editing.

ACKNOWLEDGEMENTS

This work was supported by the Chinese Academy of Sciences.

Mugui Wang[1*], Yanfei Mao[1], Yuming Lu[1], Zhidan Wang[1], Xiaoping Tao[1] and Jian-Kang Zhu[1,2*]

1. Shanghai Center for Plant Stress Biology and Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai 201602, China
2. Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, IN 47907, USA

*Correspondences: Mugui Wang (mgwang@sibs.ac.cn); Jian-Kang Zhu (jkzhu@sibs.ac.cn, Dr. Zhu is fully responsible for the distribution of all materials associated with this article)

doi: 10.1111/jipb.12667

Edited by: Li-Jia Qu, Peking University, China

Received Mar. 23, 2018; **Accepted** May 10, 2018; **Online on** May 15, 2018

FA: Free Access

AUTHOR CONTRIBUTIONS

J.Z. and M.W. conceived the project, M.W., Y.M. and Y.L. designed the experiments. M.W. generated the constructs. X.T. performed the rice transformations. M.W. and Z.W. analyzed the transformation results. M.W. drafted and J.Z. revised the manuscript.

REFERENCES

Čermák T, Curtin SJ, Gil-Humanes J, Čegan R, Kono TJY, Konečná E, Belanto JJ, Starker CG, Mathre JW, Greenstein RL, Voytas DF (2017) A multipurpose toolkit to enable advanced genome engineering in plants. **Plant Cell** 29: 1196–1217

Cody WB, Scholthof HB, Mirkov TE (2017) Multiplexed gene editing and protein overexpression using a tobacco mosaic virus viral vector. **Plant Physiol** 175: 23–35

(C) T-DNA constructs of TCTU and SSTU systems for multiplex gene editing with SpCas9 and the resulting mutagenesis efficiency at rice CYP81A family genes. Mutation rate was calculated as the number of mutants divided by the total number of plantlets genotyped for each target site. DR, direct repeat; HH, hammerhead ribozyme; HDV, hepatitis delta virus ribozyme.

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

Gao Y, Zhao Y (2014) Self-processing of ribozyme-flanked RNAs into guide RNAs *in vitro* and *in vivo* for CRISPR-mediated genome editing. **J Integr Plant Biol** 56: 343–349

Mikami M, Toki S, Endo M (2017) In planta processing of the SpCas9-gRNA complex. **Plant Cell Physiol** 58: 1857–1867

Ma X, Zhang Q, Zhu Q, Liu W, Chen Y, Qiu R, Wang B, Yang Z, Li H, Lin Y, Xie Y, Shen R, Chen S, Wang Z, Chen Y, Guo J, Chen L, Zhao X, Dong Z, Liu YG (2015) A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. **Mol Plant** 8: 1274–1284

Świat MA, Dashko S, den Ridder M, Wijsman M, van der Oost J, Daran JM, Daran-Lapujade P (2017) FnCpf1: A novel and efficient genome editing tool for *Saccharomyces cerevisiae*. **Nucleic Acids Res** 45: 12585–12598

Tang X, Zheng X, Qi Y, Zhang D, Cheng Y, Tang A, Voytas DF, Zhang Y (2016) A single transcript CRISPR-Cas9 system for efficient genome editing in plants. **Mol Plant** 9: 1088–1091

Tang X, Lowder LG, Zhang T, Malzahn AA, Zheng X, Voytas DF, Zhong Z, Chen Y, Ren Q, Li Q, Kirkland ER, Zhang Y, Qi Y (2017) A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. **Nat Plants** 3: 17018

Tu M, Lin L, Cheng Y, He X, Sun H, Xie H, Fu J, Liu Ch, Li J, Chen D, Xi H, Xue D, Liu Q, Kirkland ER, Zhang Y, Qi Y (2017) A 'new lease of life': FnCpf1 possesses DNA cleavage activity for genome editing in human cells. **Nucleic Acids Res** 45: 11295–11304

Wang M, Mao Y, Lu Y, Tao X, Zhu JK (2017) Multiplex gene editing in rice using the CRISPR-Cpf1 system. **Mol Plant** 10: 1011–1013

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

Yu Z, Chen Q, Chen W, Zhang X, Mei F, Zhang P, Zhao M, Wang X, Shi N, Jackson S, Hong Y (2018) Multigene editing via CRISPR/Cas9 guided by a single-sgRNA seed in *Arabidopsis*. **J Integr Plant Biol** 5: 376–381

Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV, Zhang F (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. **Cell** 163: 759–771

Zhang Z, Mao Y, Si H, Liu W, Botella JR (2016) A multiplex CRISPR/Cas9 platform for fast and efficient editing of multiple genes in *Arabidopsis*. **Plant Cell Rep** 35: 1519–1533

Zhong G, Wang H, Li Y, Tran MH, Farzan M (2017) Cpf1 proteins excise CRISPR RNAs from mRNA transcripts in mammalian cells. **Nat Chem Biol** 13: 839–841

Zhong Z, Zhang Y, You Q, Tang X, Ren Q, Liu S, Yang L, Wang Y, Liu X, Liu B, Zhang T, Zheng X, Le Y, Zhang Y, Qi Y (2018) Plant genome editing using FnCpf1 and LbCpf1 nucleases at redefined and altered PAM sites. **Mol Plant** doi: 10.1016/j.molp.2018.03.008

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article: http://onlinelibrary.wiley.com/doi/10.1111/jipb.12667/suppinfo

Figure S1. Configuration of TCTU-1 and TCTU-2 systems for editing rice LEA_1 and LEA_2 family genes with FnCpf1

DR, direct repeat.

Figure S2. Different configurations of crRNA expression cassette for LbCpf1 to edit the *OsPDS-2* target site and the resultant mutation rate

The data in the first configuration was derived from our previous study (Wang et al. 2017). Bi, bi-allele; Chi, Chimera; DR, direct repeat; HDV, hepatitis delta virus ribozyme; He, heterozygote; HH, hammerhead ribozyme; WT, wild type.

Figure S3. Representative mutations in targets 4,353, 5,071, 1,258 and 6,262 derived from the TCTU system of FnCpf1

The genotypes, detail on mutations and original sequencing chromatograms for each target site are listed. PAM-guide sequence is marked in grey and the PAM motif (TTN or TTTV) is marked in bold. Each dashed line represents a deleted nucleotide and the number after '-' represents the number of bases that have been deleted. Bi: bi-allele, He: heterozygous, WT: wild type.

Figure S4. Representative mutations in targets 4,998, 2,387 and 2,191 derived from the TCTU system of FnCpf1

The genotypes, detail on mutations and original sequencing chromatograms for each target site are listed. PAM-guide sequence is marked in grey and the PAM motif (TTN or TTTV) is marked in bold. Each dashed line represents a deleted nucleotide and the number after '-' represents the number of bases that have been deleted. Bi: bi-allele, He: heterozygous, WT: wild type.

Figure S5. Representative mutations in targets 4,353, 5,071, 1,258 and 6,262 derived from the SSTU system of FnCpf1

The genotypes, detail on mutations and original sequencing chromatograms for each target site are listed. PAM-guide sequence is marked in grey and the PAM motif (TTN or TTTV) is marked in bold. Each dashed line represents a deleted nucleotide and the number after '-' represents the number of bases that have been deleted. Bi: bi-allele, He: heterozygous, WT: wild type.

Figure S6. Representative mutations in targets 4,998, 2,387 and 2,191 derived from the SSTU system of FnCpf1

The genotypes, detail on mutations and original sequencing chromatograms for each target site are listed. PAM-guide sequence is marked in grey and the PAM motif (TTN or TTTV) is marked in bold. Each dashed line represents a deleted nucleotide and the number after '-' represents the number of bases that have been deleted. Bi, bi-allele; He, heterozygous; WT, wild type.

Figure S7. Representative mutations in targets 5,070, 4,487, 4,528-2 and 2,657 derived from the TCTU system of LbCpf1

The genotypes, detail on mutations and original sequencing chromatograms for each target site are listed. PAM-guide sequence is marked in grey and the PAM motif (TTTV) is marked in bold. Each dashed line represents a deleted nucleotide and the number after '-' represents the number of bases that have been deleted. Bi, bi-allele; He, heterozygous; WT, wild type.

Figure S8. Representative mutations in targets 2,675, 2,676, 2,678 and 2,679 derived from the TCTU system of LbCpf1

The genotypes, detail on mutations and original sequencing chromatograms for each target site are listed. PAM-guide sequence is marked in grey and the PAM motif (TTTV) is marked in bold. Each dashed line represents a deleted nucleotide and the number after '-' represents the number of bases that have been deleted. Bi, bi-allele; He, heterozygous; Ho, homozygous; WT, wild type.

Figure S9. Representative mutations in targets 5,070, 4,487, 4,528-2 and 2,657 derived from the SSTU system of LbCpf1

The genotypes, detail on mutations and original sequencing chromatograms for each target site are listed. PAM-guide sequence is marked in grey and the PAM motif (TTTV) is marked in bold. Each dashed line represents a deleted nucleotide and the number after '-' represents the number of bases that have been deleted. Bi, bi-allele; He, heterozygous; WT, wild type.

Figure S10. Representative mutations in targets 2,675, 2,676, 2,678 and 2,679 derived from the SSTU system of LbCpf1

The genotypes, detail on mutations and original sequencing chromatograms for each target site are listed. PAM-guide sequence is marked in grey and the PAM motif (TTTV) is marked in bold. Each dashed line represents a deleted nucleotide and the number after '-' represents the number of bases that have been deleted. Bi, bi-allele; He, heterozygous; WT, wild type.

Figure S11. Representative mutations derived from the TCTU-3 system of Cas9

The genotypes, detail on mutations and original sequencing chromatograms for each target site are listed. Guide-PAM sequence is marked in grey and the PAM motif (NGG) is marked in bold. Each dashed line represents a deleted nucleotide and the number after '-' and '+' represents the number of bases that have been deleted and inserted, respectively. Bi, bi-allele; Ho, homozygous.

Figure S12. Representative mutations derived from the SSTU system of Cas9

The genotypes, detail on mutations and original sequencing chromatograms for each target site are listed. Guide-PAM sequence is marked in grey and the PAM motif (NGG) is marked in bold. Each dashed line represents a deleted nucleotide and the number after '-' represents the number of bases that have been deleted. Bi, bi-allele; He, heterozygous; Ho, homozygous; WT, wild type.

Table S1. Summary of the multiplex gene editing in rice LEA_1 and LEA_2 family genes by TCTU-1 and TCTU-2 systems with FnCpf1

Table S2. Summary of the multiplex gene editing in rice LEA_1 and LEA_2 family genes by TCTU and SSTU systems with FnCpf1

Table S3. Summary of the multiplex gene editing in rice LEA_Dehydrin family genes by TCTU and SSTU systems with LbCpf1

Table S4. Mutants derived from the TCTU system of FnCpf1

Table S5. Mutants derived from the SSTU system of FnCpf1

Table S6. Mutants derived from the TCTU system of LbCpf1

Table S7. Mutants derived from the SSTU system of LbCpf1

Table S8. Summary of the multiplex gene editing in rice CYP81A family genes by TCTU-3 and SSTU systems with Cas9

Table S9. Mutants derived from the TCTU-3 system of Cas9

Table S10. Mutants derived from the SSTU system of Cas9

Table S11. PCR primers used to amplify the target sites in rice plantlets

Table S12. Primers used to sequence the target PCR product.

Supplementary Materials and Methods