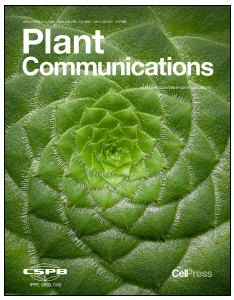
Journal Pre-proof



IsDge10 is a hypercompact TnpB nuclease that confers efficient genome editing in rice

Rui Zhang, Xu Tang, Yao He, Yangcun Li, Wei Wang, Yawei Wang, Danning Wang, Xuelian Zheng, Yiping Qi, Yong Zhang

PII: S2590-3462(24)00422-X

DOI: https://doi.org/10.1016/j.xplc.2024.101068

Reference: XPLC 101068

To appear in: *PLANT COMMUNICATIONS*

Received Date: 9 June 2024

Revised Date: 15 July 2024

Accepted Date: 19 August 2024

Please cite this article as: Zhang, R., Tang, X., He, Y., Li, Y., Wang, W., Wang, Y., Wang, D., Zheng, X., Qi, Y., Zhang, Y., IsDge10 is a hypercompact TnpB nuclease that confers efficient genome editing in rice, *PLANT COMMUNICATIONS* (2024), doi: https://doi.org/10.1016/j.xplc.2024.101068.

Correspondence

IsDge10 is a hypercompact TnpB nuclease that confers efficient genome editing in rice

Rui Zhang[1, 2], Xu Tang[2, 3], Yao He[2], Yangcun Li[2], Wei Wang[1], Yawei Wang[1], Danning Wang[1], Xuelian Zheng[2], Yiping Qi[4, 5*], Yong Zhang[2*]

[1]Department of Biotechnology, School of Life Sciences and Technology, Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu 610054, China; [2]Integrative Science Center of Germplasm Creation in Western China (Chongqing) Science City, Chongqing Key Laboratory of Tree Germplasm Innovation and Utilization, School of Life Sciences, Southwest University, 400715 Chongqing, China; [3]State Key Lab of Rice Biology, China National Rice Research Institute, Hangzhou 311400, Zhejiang, China; [4]Department of Plant Science and Landscape Architecture, University of Maryland, College Park, Maryland 20742, USA; [5]Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, Maryland 20850;

Running title: Hypercompact IsDge10 enables genome editing in rice

***Corresponding authors:**

Yong Zhang, Integrative Science Center of Germplasm Creation in Western China (Chongqing) Science City, Chongqing Key Laboratory of Tree Germplasm Innovation and Utilization, School of Life Sciences, Southwest University, Chongqing 400715, China; Email: zhangyong916@swu.edu.cn

Yiping Qi, Department of Plant Science and Landscape Architecture, University of Maryland, College Park, MD 20742, USA; Email: yiping@umd.edu

Dear Editor,

Cas9 and Cas12a have been widely applied in genome engineering in both plant and human cells (Tang and Zhang, 2023). However, the large size of these proteins prevents their delivery into cells via viral vectors. As the hypothetical ancestors of Cas9 and Cas12a, TnpB and IscB have been reported as RNA-guided DNA endonucleases for genome editing in human cells (Han et al., 2023; Xiang et al., 2023). More recently, Fanzor has been reported as an eukaryotic RNA-guided endonuclease to possess genome editing activity in human cells (Saito et al., 2023). These nucleases (e.g., ~390 amino acids for IsDge10) are much smaller than Cas9 and Cas12a (e.g., ~1300 amino acids for SpCas9). However, it was not clear whether TnpB, IscB, or Fanzor could be harnessed for genome engineering in plants. Here, we evaluated a series of members from TnpB, IscB, and Fanzor families, and successfully developed a miniature plant genome editor based on IsDge10, a TnpB nuclease from *Deinococcus geothermalis.*

First, we selected six different nucleases from these three small nucleases families, including IsDge10, IsAam1, enIscB, and SpuFz1 (**Figure 1A**) (Han *et al.*, 2023; Saito *et al.*, 2023; Xiang *et al.*, 2023), and we optimized their codons for rice expression. Next, we used the RNA polymerase II promoter, ZmUbi1, to drive the expression of the nuclease gene and the OsU6-2 promoter to drive the expression of their respective guide RNAs (**Figure 1B**). Then, we developed a dual-fluorescence reporter system that simultaneously expresses GFP and mCherry. The green fluorescence from GFP serves as a normalization standard, while the red fluorescence from mCherry can be perturbated by targeted mutagenesis by any of these seven nuclease systems (**Supplementary Figure 1A**). The reporter system was co-transfected with the nuclease system into rice protoplasts, allowing us to preliminarily assess the editing capability of our constructed nuclease systems in rice protoplasts. In our design, the mCherry gene was targeted at one site with the transposon-associated motif (TAM) by each corresponding nuclease (**Figure 1A and Supplementary Table 1**). The

results showed that IsDge10, IsAam1, enIscB, and SpuFz1 exhibited detectable editing activity in rice protoplasts, based on the reduction of mCherry to GFP ratios, although their editing activity appeared to lower than Cas9 (**Supplementary Figure 1B and 1C**). Therefore, we chose these four nucleases for further experiments.

To test the editing efficiency of these four systems at endogenous sites in rice, we chose seven sites per nuclease system and assessed the targeted mutagenesis outcomes in rice protoplasts with next-generation sequencing (NGS) of PCR amplicons (**Figure 1A and Supplementary Table 1**). Our data showed that IsDge10 exhibited 2.20%-15.04% editing efficiency across seven target sites. In contrast, enIscB exhibited 2.05%-8.27% editing efficiency at five of seven sites; IsAam1 exhibited 2.36%-4.65% editing efficiency at two of seven sites; SpuFz1 showed no editing efficiency (**Figure 1C and Supplementary Figure 2**). These data suggest that IsDge10 is superior to the enIscB, IsAam1 and SpuFz1 systems for genome editing in rice. IsDge10 predominantly generated deletions and the deletion size ranged from 6 to 10 bp, with the deletion position located 13 to 23 bp away from the TAM (**Figure 1D, 1E, and Supplementary Figure 3-5**). This cleavage pattern of IsDge10 mirrors those of Cas12 nucleases, is different from that of enIscB, which cleaves at the TAM proximal region, consistent with enIscB's evolutionary relevance to the Cas9 group (**Supplementary Figure 6, 7**). These targeted mutagenesis characteristics suggest that IsDge10, like Cas12 nucleases, generates off-set DNA double strand breaks distal to the TAM sites.

To assess the specificity of IsDge10, we focused on a high-activity target site, site 01, and designed a series of protospacers introducing two adjacent mutations at different positions (**Supplementary Table1 and Supplementary Figure 2A**). The protoplast results showed that permutations of every two nucleotides at positions 1-14 bp of the protospacer completely abolished the editing activity of IsDge10 while mutations at the 15-16 bp of the protospacer

caused ~50% reduction of the editing frequency (**Figure 1F**). In contrast, mutations at positions 17-20 bp of the protospacer did not significantly affect the editing activity of IsDge10 (**Figure 1F**). This experiment defined the core length of a functional spacer for IsDge10 and showed that IsDge10 is a highly specific nuclease.

We next tested whether IsDge10 could generate edits in stable rice lines. The T-DNA constructs targeting the same seven sites were used for stable rice transformation. The results showed that mutated T0 generation plants could be obtained at all seven sites, with mutation efficiency ranging from 4.2% to 25% (**Figure 1G**). Analysis of mutant plants showed that the mutations were mainly deletions of 5 bp to 10 bp or longer (**Supplementary Figure 8**), which was consistent with the editing profile observed in rice protoplasts (**Figure 1D**). Interestingly, we only recovered monoallelic mutations in these rice lines (**Supplementary Figure 9**). Assuming biallelic knockout of these genes is not lethal, these data suggest that IsDge10 may be further improved to render biallelic editing in rice.

To establish a multiplexed IsDge10 genome editing system, we used the dual RNA polymerase II promoter expression systems previously employed for Cas12a, Cas12b, and Cas12j2, and their guide RNAs (Liu et al., 2022; Ming et al., 2020; Tang et al., 2017; Tang et al., 2019; Zheng et al., 2023; Zhou et al., 2023). The IsDge10 protein was expressed under the ZmUbi1 promoter, and the seven guide RNAs were expressed with the OsUbi1 promoter and processed by the HH-HDV dual ribozyme system to form mature “guide RNA-ωRNA” (**Figure 1H**). Interestingly, this multiplexed construct demonstrated higher editing efficiencies at all seven target sites compared to the OsU6-2 promoter driving the expression of their respective guide RNAs (**Figure 1C and Supplementary Figure 2A**), with efficiencies ranging from approximately 4.3% to 18.2% in rice protoplasts (**Figure 1I**). Deletion sizes were typically 6-10 bp and deletion position located about 13-23 bp away from the TAM (**Figure 1J-K**

117 **and Supplementary Figure 10, 11**). These results demonstrate that IsDge10
118 can simultaneously edit multiple sites, based on this robust dual Pol II promoter
119 system.

120 In summary, our study demonstrates that IsDge10 is a novel and small
121 transposon-associated TnpB nuclease for genome editing in rice. Compared to
122 other compact nucleases tested, IsDge10 has robust genome editing activity in
123 rice and a simple TTAT TAM requirement. Although our data suggest the current
124 IsDge10 system is not as efficient as the commonly used Cas9 and Cas12a
125 systems, this study paves the road for further improvement with protein
126 engineering and evolution. As one of the smallest nucleases functional in plants,
127 IsDge10 holds great potential in various aspects such as multi-nuclease
128 combination editing, fusion with different effectors to develop transcriptional and
129 epigenetic regulation toolkits, as well as in viral delivery in plant genome
130 engineering.

SUPPLEMENTAL INFORMATION

Supplemental information is available in the online version of this article.

FUNDING

This research was supported by the Biological Breeding-Major Projects (2023ZD04076) to X.T. and Y.Z., the National Natural Science Foundation of China (award no. 32270433, 32101205, and 32072045) to Y.Z., X.L.Z. and X.T. It is also supported by the NSF Plant Genome Research Program (IOS-2029889 and IOS-2132693) to Y.Q.

AUTHORS' CONTRIBUTIONS

Y.Z. proposed the project and designed the experiments. R.Z., X.T., and Y.H. generated all the constructs. R.Z. did rice protoplast transformation and analyzed the mutation frequencies in protoplasts. R.Z., W.W., Y.W., D.W., and X.Z. conducted rice stable transformation. Y.L assisted in revising the manuscript. Y.Z., Y.Q., and R.Z. analyzed the data and wrote the paper with input from other authors. All authors read and approved the final version of the manuscript.

DECLARATION OF INTERESTS

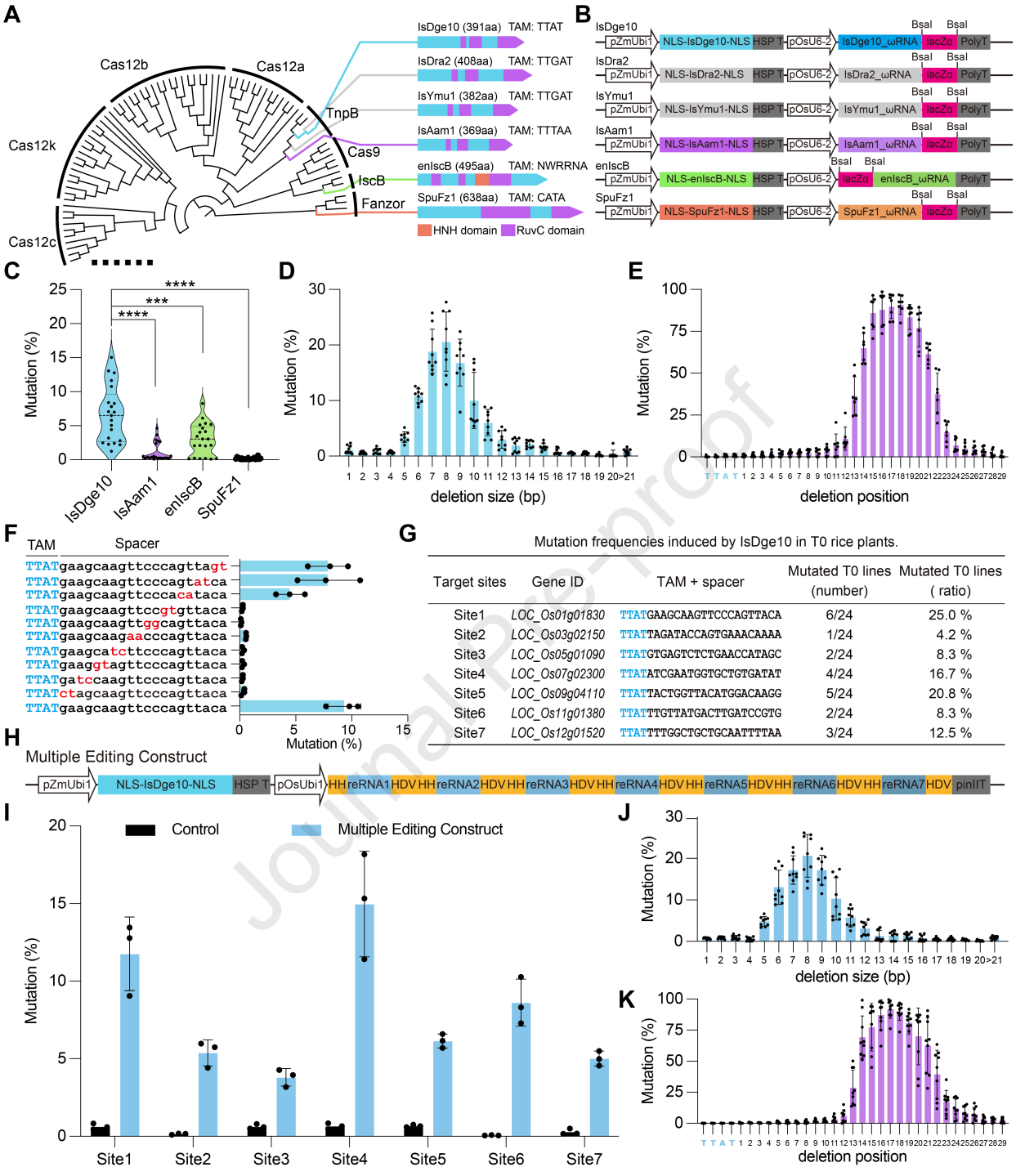
The authors declare no competing interests.

References

- **Han, D., Xiao, Q., Wang, Y., Zhang, H., Dong, X., Li, G., Kong, X., Wang, S., Song, J., Zhang, W., et al.** (2023). Development of miniature base editors using engineered IscB nickase. Nat Methods 10.1038/s41592-023-01898-9.
- **Liu, S., Sretenovic, S., Fan, T., Cheng, Y., Li, G., Qi, A., Tang, X., Xu, Y., Guo, W., Zhong, Z., et al.** (2022). Hypercompact CRISPR-Cas12j2 (CasPhi) enables genome editing, gene activation, and epigenome editing in plants. Plant Commun **3**:100453. 10.1016/j.xplc.2022.100453.
- **Ming, M., Ren, Q., Pan, C., He, Y., Zhang, Y., Liu, S., Zhong, Z., Wang, J., Malzahn, A.A., Wu, J., et al.** (2020). CRISPR-Cas12b enables efficient plant genome engineering. Nat Plants **6**:202-208. 10.1038/s41477-020-0614-6.
- **Saito, M., Xu, P., Faure, G., Maguire, S., Kannan, S., Altae-Tran, H., Vo, S., Desimone, A., Macrae, R.K., and Zhang, F.** (2023). Fanzor is a eukaryotic programmable RNA-guided endonuclease. Nature 10.1038/s41586-023-06356-2.
- **Tang, X., and Zhang, Y.** (2023). Beyond knockouts: fine-tuning regulation of gene expression in plants with CRISPR-Cas-based promoter editing. New Phytol **239**:868-874. 10.1111/nph.19020.
- **Tang, X., Lowder, L.G., Zhang, T., Malzahn, A.A., Zheng, X., Voytas, D.F., Zhong, Z., Chen, Y., Ren, Q., Li, Q., et al.** (2017). A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. Nat Plants **3**:17018. 10.1038/nplants.2017.18.
- **Tang, X., Ren, Q., Yang, L., Bao, Y., Zhong, Z., He, Y., Liu, S., Qi, C., Liu, B., Wang, Y., et al.** (2019). Single transcript unit CRISPR 2.0 systems for robust Cas9 and Cas12a mediated plant genome editing. Plant Biotechnol J **17**:1431-1445. 10.1111/pbi.13068.
- **Xiang, G., Li, Y., Sun, J., Huo, Y., Cao, S., Cao, Y., Guo, Y., Yang, L., Cai, Y., Zhang, Y.E., et al.** (2023). Evolutionary mining and functional characterization of TnpB nucleases identify efficient miniature genome editors. Nat Biotechnol 10.1038/s41587-023-01857-x.
- **Zheng, X., Zhang, S., Liang, Y., Zhang, R., Liu, L., Qin, P., Zhang, Z., Wang, Y., Zhou, J., Tang, X., et al.** (2023). Loss-function mutants of OsCKX gene family based on CRISPR-Cas systems revealed their diversified roles in rice. Plant Genome **16**:e20283. 10.1002/tpg2.20283.
- **Zhou, J., Liu, G., Zhao, Y., Zhang, R., Tang, X., Li, L., Jia, X., Guo, Y., Wu, Y., Han, Y., et al.** (2023). An efficient CRISPR-Cas12a promoter editing system for crop improvement. Nat Plants **9**:588-604. 10.1038/s41477-023-01384-2.

Figure legend

Figure 1. Development of the IsDge10 genome editing system in rice. (A) Phylogenetic diagram illustrating the evolutionary relationships among TnpB, Cas9, Cas12, IscB, and Fanzor. The selected nucleases IsDge10, IsDra2, IsYmu1, IsAam1, enIscB, and SpuFz1 are highlighted and connected to their corresponding structural diagrams, showing their sizes and domains (HNH and RuvC). (B) Schematics of the IsDge10, IsDra2, IsYmu1, IsAam1, enIscB and SpuFz1 constructs for genome editing in rice. (C) Comparing the mutation rates of IsDge10, IaAam1, enIscB and SpuFz1 systems in rice protoplasts. (D) Deletion size profile for three representative target sites in rice. (E) Deletion position profile for three representative target sites in rice. (F) Assessment of targeting specificity using mismatched guide RNAs at a representative target site in rice protoplasts. (G) Genome editing efficiency of IsDge10 in stable rice lines at seven target sites. (H) Schematics of the dual RNA polymerase II promoter-based and multiplexed IsDge10 system for genome editing in rice. (I) Multiplexed editing of seven target sites in rice protoplasts. (J) Deletion size profile for three representative multiplexed target sites in rice. (K) Deletion position profile for three representative multiplexed target sites in rice. Each dot represents a biological replicate. Data are presented as mean values +/- SD. Data were analyzed using two-tailed unpaired *t*-Test. ***, $P<0.001$; ****, $P<0.0001$. Solid line, median; dash line, quartiles.



Target sites	Gene ID	TAM + spacer	Mutated T0 lines (number)	Mutated T0 lines (ratio)
Site1	*LOC_Os01g01830*	TTATGAAGCAAGTTCCCAGTTACA	6/24	25.0 %
Site2	*LOC_Os03g02150*	TTATTAGATACCAGTGAAACAAAA	1/24	4.2 %
Site3	*LOC_Os05g01090*	TTATGTGAGTCTCTGAACCATAGC	2/24	8.3 %
Site4	*LOC_Os07g02300*	TTATATCGAATGGTGCTGTGATAT	4/24	16.7 %
Site5	*LOC_Os09g04110*	TTATTACTGGTTACATGGACAAGG	5/24	20.8 %
Site6	*LOC_Os11g01380*	TTATTTGTTATGACTTGATCCGTG	2/24	8.3 %
Site7	*LOC_Os12g01520*	TTATTTTGGCTGCTGCAATTTTAA	3/24	12.5 %