# An Intein-Mediated Split−nCas9 System for Base Editing in Plants

Guoliang Yuan, Haiwei Lu, Kuntal De, Md Mahmudul Hassan, Yang Liu, Yi Li, Wellington Muchero, Paul E. Abraham, Gerald A. Tuskan,* and Xiaohan Yang*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** Virus-assisted delivery of the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) system represents a promising approach for editing plant genomes. Among the CRISPR/Cas systems, CRISPR/Cas9 is most widely used; however, to pack the relatively large size of the CRISPR/Cas9 system into viral vectors with confined packaging capacity is challenging. To address this technical challenge, we developed a strategy based on split inteins that splits the required CRISPR/Cas9 components across a dual-vector system. The CRISPR/Cas reassembles into an active form following co-infection to achieve targeted genome editing in plant cells. An intein-mediated split system was adapted and optimized in plant cells by a successful demonstration of split-eYGFPuv expression. Using a plant-based biosensor, we demonstrated for the first time that the split-nCas9 can induce efficient base editing in plant cells. We identified several split sites for future biodesign strategies. Overall, this strategy provides new opportunities to bridge different CRISPR/Cas9 tools including base editor, prime editor, and CRISPR activation with virus-mediated gene editing.



**KEYWORDS:** *CRISPR/Cas9, base editing, split−SpnCas9, eYGFPuv, biosensor, transient gene expression*
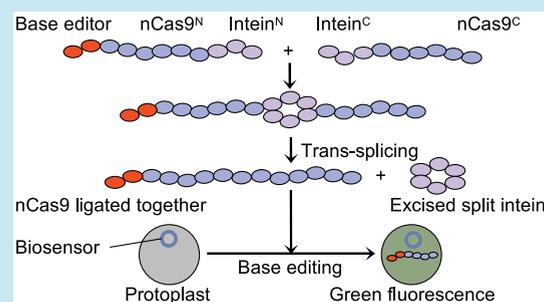
## INTRODUCTION

CRISPR/Cas-based technology has revolutionized plant genome editing.[1] However, there are still essential limitations that currently impede the commercial applications of this technology in agriculture, particularly the presence of transgenes. Virus/nanoparticle-mediated deliveries of CRISPR/Cas systems are promising methods to create transgene-free targeted mutants without requiring the lengthy tissue culture process.[2,3] RNA viruses-assisted delivery of CRISPR/Cas systems has been successfully demonstrated in plants as a method that does not involve tissue culture; however, only until recently, the positive-strand RNA viruses, such as Tobacco rattle virus (TRV) or Barley stripe mosaic virus, have mainly been used for sgRNA delivery due to their limited cargo capacity.[4,5]

An effective approach to reduce the size of a gene transcription unit is using a split-protein with or without the help of dimerization domains. An intein-mediated split-*Sreptococcus pyogenes* Cas9 (SpCas9) system was recently demonstrated in human cells, whereby the SpCas9 nuclease protein coding system was functionally split into two inactivate fragments across a dual-vector system and delivered, and its activity was reconstituted efficiently in cells via coexpression.[6] Two types of "flexible linker regions"-based split *Staphylococcus aureus* Cas9 (SaCas9) (430N/431C and 739N/740C) have been shown to exhibit genome editing activity in human cells.[7] Recently, these two types of split-SaCas9 have also been applied in plants to induce targeted mutagenesis for transgene-

free genome editing.[8] To date, the reported split sites of Cas9 are still limited, and less is known about how to effectively split the most commonly used SpCas9 for plant genome editing. Here, we developed an intein-mediated split SpCas9 nickase (SpnCas9, D10A) for base editing in plant protoplasts that functions with high efficiency and comparable performance to wild-type full-length SpnCas9.

## RESULTS AND DISCUSSION

Protein splicing elements (called "inteins") allow the coding sequence of a target protein to be split into two inactive fragments, which can then be reconstituted post-translationally (Figure 1a). We used one split intein, derived from *Npu*DnaE[9] for splitting SpnCas9. *Npu*DnaE intein was codon optimized for *Arabidopsis* to improve gene expression and translational efficiency. The coding sequence of a target gene was split into an N-terminal fragment (Gene^N) and a C-terminal fragment (Gene^C), which were then cloned upstream of an N-terminal fragment of the *Npu*DnaE intein (Int^N) and downstream of a
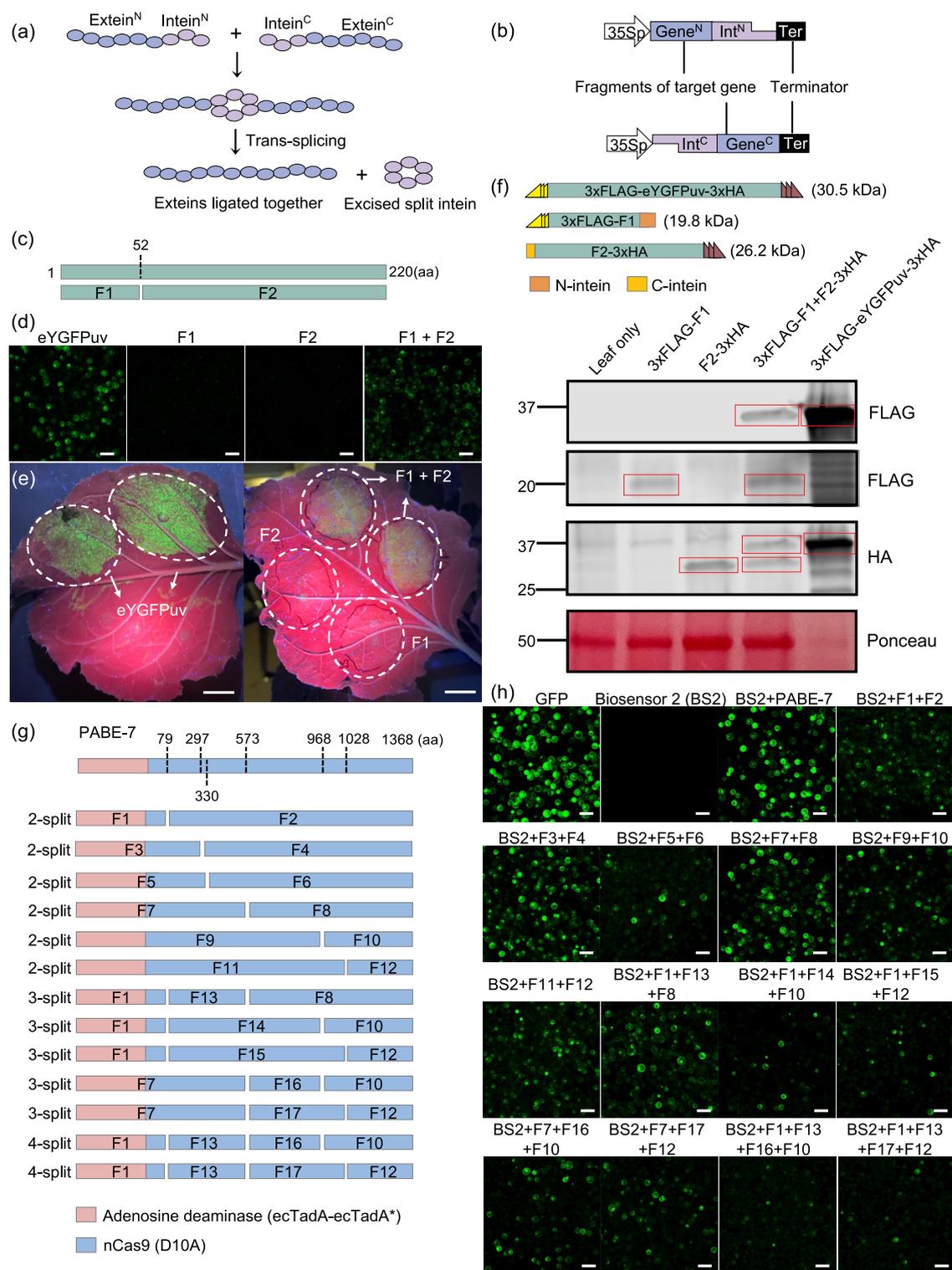
**Figure 1.** The *Npu*DnaE intein-mediated split−SpnCas9 for base editing in plant system. (a) Trans-splicing mechanism reaction by split inteins. (b) Illustration of vector design of a target gene. (c) Identification of potential split site of eYGFPuv. (d) Transient expression of split-eYGFPuv in *Arabidopsis* protoplasts. Scale bar, 100 $\mu$m. (e) Transient expression of split-eYGFPuv in *N. benthamiana*. Scale bar, 1 cm. (f) Western blot analysis of trans-splicing of the eYGFPuv protein into two fragments F1 and F2. A full-length eYGFPuv is N-terminally tagged with 3xFLAG epitope and C-terminally tagged with 3xHA epitope. The fragment F1 is N-terminally tagged with 3xFLAG epitope while the fragment F2 is C-terminally tagged with HA epitope. Western blot was performed with the proteins extracted from tobacco leaves, which were either transfected with the plasmids containing the full-length eYGFPuv and split eYGFPuv fragments, respectively, or cotransfected with an F1-containing plasmid and an F2-containing plasmid, using the antibodies indicated on the right. The bands corresponding to the F1-fragment, F2-fragment, or reconstituted eYGFPuv proteins are indicated with a red box, with 3xFLAG-eYGFPuv-3xHA as a positive control. Ponceau serves as the equal-loading control. Less positive control was loaded to prevent a smeared band due to its high protein expression. (g) Identification of potential split sites of SpnCas9. (h) Transient expression of split-SpnCas9 in *Arabidopsis* protoplasts. Scale bar, 100 $\mu$m.
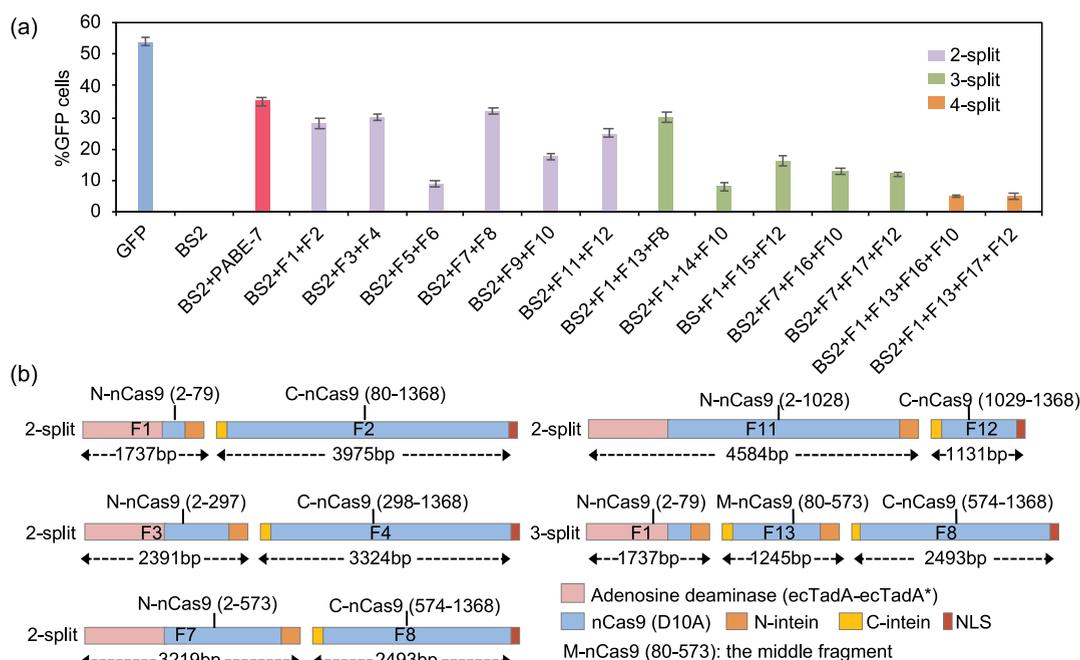
**Figure 2.** The identification of valid split sites for split−SpnCas9. (a) Statistical analysis of GFP-positive cells with different split-SpnCas9 components. All data are presented as the mean ± SE ($n$ = 5 independent scopes). (b) Different SpnCas9 fragments with high editing efficiency. NLS, nuclear localization signals.

C-terminal fragment of the *Npu*DnaE intein (Int$^C$), respectively, into two vectors (Figure 1b).

An eYGFPuv reporter[10] was selected to test the efficacy of the split system. Given that the obligatory cysteine residue on the C-extein junction and a residue on the N-extein junction promote substantial trans-splicing activities,[9] one potential split site T52:C53 was identified to split the eYGFPuv into two fragments, F1 and F2 (Figure 1c). Then two plasmids containing F1 and F2, respectively, were cotransformed into *Arabidopsis* protoplasts with a positive control 35Sp-eYGFPuv. Bright green fluorescence was observed under a confocal microscope in both the positive control and the protoplasts of F1 and F2 cotransformation, though the fluorescence in the latter was relatively weaker, whereas no fluorescence was detected in the protoplasts containing F1/F2 plasmids alone (negative controls) (Figure 1d and Supplementary Figure 1a). Meanwhile, we also coexpressed F1′ and F2′ without *Npu*DnaE intein in protoplasts and no GFP signals were detected (Supplementary Figure 2). Together, it supports that the functional eYGFPuv protein was restored by *Npu*DnaE intein not self-association of split fluorescent proteins. Furthermore, we tested split-eYGFPuv using *Agrobacterium*-mediated leaf infiltration in *Nicotiana benthamiana*. Similarly, green fluorescence was observed under UV light in the positive control and the leaf area co-infiltrated with F1 and F2 plasmids but not in the negative controls (Figure 1e). To directly observe protein splicing, we conducted Western blot analysis of protein trans-splicing between the fragment F1 (N-terminally tagged with 3xFLAG-epitope) and the fragment F2 (C-terminally tagged with 3xHA-epitope) (Figure 1f and Supplementary Figure 3). As expected, F1 and F2 with matching N- and C-inteins supported reconstitution of the full-length eYGFPuv. These results indicate that the *Npu*DnaE intein-based split system works efficiently in plant systems.

To split SpnCas9, we identified two native split sites (I79:C80 and E573:C574) and four artificial split sites

(S297^C, Q330^C, K968^C, and E1028^C) by inserting a cysteine on the C-extein junction (Figure 1g). The artificial split sites were chosen based on the length of split fragments. To detect base editing activities in plant systems, we previously developed biosensor 2 (BS2) that is composed of a *GFP* mutant harboring a premature termination codon (PTC) and a single guide RNA (sgRNA) targeting the PTC. The *GFP* mutant can be rescued by a plant adenine base editor PABE-7 under the guidance of the sgRNA, leading to the generation of green fluorescence.[11] Here, the efficacy of the split SpnCas9 system was examined in *Arabidopsis* protoplasts by the cotransfection with BS2. Bright green fluorescence was observed in both the positive control (35Sp:GFP) and the protoplasts cotransfected with a PABE-7 and BS2 but not those transformed with BS2 alone, indicating that BS2 detected the base editing activity successfully in protoplasts (Figure 1h and Supplementary Figure 1b). In contrast, strong green fluorescence was detected in the protoplasts coexpressing BS2 and 2-split SpnCas9 fragments, indicating that the intein-mediated split SpnCas9 induced active base editing in plant cells. Moreover, we examined 3-split and 4-split SpnCas9 systems containing three and four fragments, respectively (Figure 1h and Supplementary Figure 1b). Interestingly, a clear GFP signal was detected in the protoplasts containing BS2 and five 3-split combinations or two 4-split combinations, while the GFP intensity was lower in comparison with those coexpressing 2-split fragments. These results indicate that the intein-reconstituted split SpnCas9 system is functional in plant-based base editing. Furthermore, ~54% and ~35% of the cells exhibited GFP signals in the positive control and the samples with BS2-PABE-7 cotransformation, respectively (Figure 2a). In the protoplasts coexpressing BS2 and different 2-split components, approximately 9−32% of the cells exhibited GFP fluorescence (Figure 2a). GFP fluorescence was also detected in about 8−30% and 5% of the cells in the three-fragment and four-fragment cotransformations, respec-

tively (Figure 2a). These results indicate that the targeting efficiency of BS2 with the intein-mediated split−SpnCas9 system is comparable to wild-type SpnCas9, especially in 2-split sites (I79:C80, S297^C, E573:C574 and E1028^C) and 3-split sites (I79:C80 and E573:C574) (Figure 2b and Supplementary Figure 4). Theoretically, this split system can be used to create a substitutable foundational domain (e.g., activator, repressor, base editor, or primer editor) of CRISPR tools to make a split system toolbox. We then examined another split site M1^C between ABE domain and SpnCas9 (Supplementary Figure 5a). However, in comparison with the positive controls, only weak GFP signals were observed in the protoplasts coexpressing intein-mediated ABE domain and SpnCas9 together with biosensor 2 (Supplementary Figure 5b). Thus, it appears that the split site I79:C80 would be a better candidate for the swapping of functional domains of CRISPR tools by now.

In summary, we showed that the NpuDnaE intein-mediated split system functions effectively in plant systems. We also demonstrated that the base editing activity of our split-intein system is comparable to wild-type full-length SpnCas9. Recently, it was reported that a split prime editor system could mediate endogenous base transversion and insertion in human cells.[12] And it has been shown that Tobacco etch virus (TEV)-based system and Potato virus X (PVX)-based vector can infect N. benthamiana simultaneously.[13] Therefore, the identification of multiple valid split sites provides more choice for split SpnCas9 and can be potentially used for the delivery of CRISPR/Cas9-based tools such as base editor, prime editor, and CRISPR interference/activation through viruses/nano-particles-mediated transformation.

## ■ MATERIALS AND METHODS

**Genetic Constructs.** To split eYGFPuv, a gBlocks Gene Fragment containing 5′-eYGFPuv and N-terminal of NpuD-naE was synthesized from Integrated DNA Technologies IDT. The fragment F1 of eYGFPuv was assembled by cloning the gBlocks Gene Fragment and a PCR-amplified relevant fragment into an eYGFPuv vector[10] through NEBuilder HiFi DNA Assembly (New England BioLabs, Catalog #E5520S). Similarly, the fragment F2 of eYGFPuv was created using a gBlocks Gene Fragment containing 3′-eYGFPuv and C-terminal of NpuDnaE. To split SpnCas9, a PCR-amplified 5′-PABE-7 fragment and a gBlocks Gene Fragment containing an N-terminal of NpuDnaE were assembled into the eYGFPuv vector, generating F1, F3, F5, F7, F9, F11, and F18. Then a PCR-amplified 3′-PABE-7 fragment and a gBlocks Gene Fragment containing a C-terminal of NpuDnaE were used to assemble F2, F4, F6, F8, F10, F12, and F19. The construction of F13 to F17 was completed by cloning a gBlocks Gene Fragment containing an N-terminal of NpuDnaE, a middle fragment of PABE-7, and a PCR-amplified C-terminal of NpuDnaE into the eYGFPuv vector. DNA sequences encoding inteins were codon optimized for Arabidopsis using the online codon optimization tool (ExpOptimizer) provided by NovoPro Bioscience (Shanghai, China). Positive plasmids were selected through colony PCR and verified by Sanger sequencing.

**Protoplast Transformation.** The isolation and transient transformation of Arabidopsis leaf mesophyll protoplasts were performed as described previously.[14] The concentration of protoplasts was adjusted to $2 \times 10^5$/mL by adding W5 Solution. A 100 $\mu$L aliquot of protoplasts was used per

reaction. To examine split-eYGFPuv, the same amount of plasmid DNA (5 $\mu$L, ~1 $\mu$g/$\mu$L) was loaded for samples eYGFP (positive control), F1, F1′, F2, and F2′. To examine split-SpnCas9, 4 $\mu$L (~1 $\mu$g/$\mu$L) of plasmid DNA was loaded for samples eYGFP and Biosensor 2 (BS2) while 3 $\mu$L (~1 $\mu$g/$\mu$L) of plasmid DNA was loaded for samples F1 to F19 and PABE-7. Each split site was tested through at least three independent experiments.

**Statistical Analysis.** The percentage of cells with GFP signal (%GFP cells) was calculated based on confocal microscope imaging. The images of five independent scopes (technical replicates) were taken under confocal microscope with 10 $\mu$L protoplasts per sample. Given that the broken cells cannot be recognized properly by cell counting software, the cell-counting in each image was conducted manually. %GFP cells (per scope) = number of GFP cells/total number of cells counted. The average %GFP cells (per sample) was presented to indicate the base editing efficiency. The number of cells in each scope varied between 50 and 300. All data are presented as the mean ± SE ($n$ = 5 independent scopes).

**Protein Extraction and Western Blot.** Transiently transformed tobacco leaves were ground using mortar and pestle. Protein extracts were generated using 25 mM Tris-HCl, pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS, and protease and phosphatase inhibitor tablets (Thermo Fisher Scientific). Extracted proteins were quantified using the Qubit (Thermo Fisher Scientific). Proteins were separated by SDS acrylamide gel electrophoresis and transferred to IMMOBILON-FL 26 PVDF membrane (Millipore) probed with the indicated antibodies and visualized by LiCor Odyssey infrared imaging system. Antibodies used: HA (#902302; 1:1000) Biolegend, M2 FLAG (1:1000) antibody from Sigma, IRDye 800CW Goat anti-Mouse IgG Secondary Antibody (LiCor), IRDye 680RD Goat anti-Rabbit IgG Secondary Antibody (LiCor).

**Microscopy Analysis.** All fluorescence images were taken using a Zeiss LSM 710 confocal microscope with a FITC (green) filter. The EGFP dye was excited within the spectral range of 493 to 584 nm. Fixed gain (master) was applied to the same batch of samples.

**Tobacco Leaf Infiltration.** Agrobacterium strain GV3101 harboring the plasmid of interest was injected into N. benthamiana leaves using a syringe without a needle as described by Li.[15]

## ■ ASSOCIATED CONTENT

ⓢ **Supporting Information**

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.1c00507.

    Supplementary Figures and vector information (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Authors

**Gerald A. Tuskan** − Biosciences Division and The Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States; Email: tuskanga@ornl.gov

**Xiaohan Yang** − Biosciences Division and The Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States; ⓘ orcid.org/0000-0001-5207-4210; Email: yangx@ornl.gov

## Authors

**Guoliang Yuan** − *Biosciences Division and The Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States;* ⬤ orcid.org/0000-0002-6562-8769

**Haiwei Lu** − *Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States;* Present Address: Department of Academic Education, Central Community College−Hastings, Hastings, NE 68902, USA

**Kuntal De** − *Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States*

**Md Mahmudul Hassan** − *Biosciences Division and The Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States; Department of Genetics and Plant Breeding, Patuakhali Science and Technology University, Dumki, Patuakhali 8602, Bangladesh*

**Yang Liu** − *Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States*

**Yi Li** − *Department of Plant Science and Landscape Architecture, University of Connecticut, Storrs, Connecticut 06269, United States*

**Wellington Muchero** − *Biosciences Division and The Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States*

**Paul E. Abraham** − *Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acssynbio.1c00507

## Author Contributions

G.Y. and X.Y conceived the research. G.Y., M.H., and Yang L. created the genetic constructs. G.Y. and H. L. conducted the protoplast transformation. K.D. and W.M. conducted the protein extraction and Western blot. G.Y. drafted the manuscript. Yi L., P.A., G.T., and all the other authors contributed to the data analysis and manuscript preparation.

## Notes

## ■ REFERENCES

(1) Hassan, M. M.; Zhang, Y.; Yuan, G.; De, K.; Chen, J.-G.; Muchero, W.; Tuskan, G. A.; Qi, Y.; Yang, X. Construct design for CRISPR/Cas-based genome editing in plants. *Trends Plant Sci.* **2021**, *26*, 1133−1152.

(2) Lv, Z.; Jiang, R.; Chen, J.; Chen, W. Nanoparticle-mediated gene transformation strategies for plant genetic engineering. *Plant Journal* **2020**, *104*, 880−891.

(3) Zhu, H.; Li, C.; Gao, C. Applications of CRISPR−Cas in agriculture and plant biotechnology. *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 661−677.

(4) Li, T.; Hu, J.; Sun, Y.; Li, B.; Zhang, D.; Li, W.; Liu, J.; Li, D.; Gao, C.; Zhang, Y.; Wang, Y. Highly efficient heritable genome editing in wheat using an RNA virus and bypassing tissue culture. *Molecular Plant* **2021**, *14*, 1787−1798.

(5) Ellison, E. E.; Nagalakshmi, U.; Gamo, M. E.; Huang, P.-j.; Dinesh-Kumar, S.; Voytas, D. F. Multiplexed heritable gene editing using RNA viruses and mobile single guide RNAs. *Nature Plants* **2020**, *6*, 620−624.

(6) Truong, D.-J. J.; Kühner, K.; Kühn, R.; Werfel, S.; Engelhardt, S.; Wurst, W.; Ortiz, O. Development of an intein-mediated split−Cas9 system for gene therapy. *Nucleic Acids Res.* **2015**, *43*, 6450−6458.

(7) Nishimasu, H.; Cong, L.; Yan, W. X.; Ran, F. A.; Zetsche, B.; Li, Y.; Kurabayashi, A.; Ishitani, R.; Zhang, F.; Nureki, O. Crystal Structure of Staphylococcus aureus Cas9. *Cell* **2015**, *162*, 1113−1126.

(8) Kaya, H.; Ishibashi, K.; Toki, S. A Split Staphylococcus aureus Cas9 as a Compact Genome-Editing Tool in Plants. *Plant and Cell Physiology* **2017**, *58*, 643−649.

(9) Jillette, N.; Du, M.; Zhu, J. J.; Cardoz, P.; Cheng, A. W. Split selectable markers. *Nat. Commun.* **2019**, *10*, 4968.

(10) Yuan, G.; Lu, H.; Tang, D.; Hassan, M. M.; Li, Y.; Chen, J.-G.; Tuskan, G. A.; Yang, X. Expanding the application of a UV-visible reporter for transient gene expression and stable transformation in plants. *Horticulture Research* **2021**, *8*, 234.

(11) Yuan, G.; Hassan, M. M.; Yao, T.; Lu, H.; Vergara, M. M.; Labbé, J. L.; Muchero, W.; Pan, C.; Chen, J.-G.; Tuskan, G. A.; Qi, Y.; Abraham, P. E.; Yang, X. Plant-Based Biosensors for Detecting CRISPR-Mediated Genome Engineering. *ACS Synth. Biol.* **2021**, *10*, 3600−3603.

(12) Zhi, S.; Chen, Y.; Wu, G.; Wen, J.; Wu, J.; Liu, Q.; Li, Y.; Kang, R.; Hu, S.; Wang, J.; Liang, P.; Huang, J. Dual-AAV delivering split prime editor system for *in vivo* genome editing. *Molecular Therapy* **2022**, *30*, 283−294.

(13) Calvache, C.; Vazquez-Vilar, M.; Selma, S.; Uranga, M.; Fernández-Del-Carmen, A.; Daròs, J.-A.; Orzáez, D. Strong and tunable anti-CRISPR/Cas activities in plants. *Plant biotechnology journal* **2022**, *20*, 399−408.

(14) Xie, M.; Muchero, W.; Bryan, A. C.; Yee, K.; Guo, H.-B.; Zhang, J.; Tschaplinski, T. J.; Singan, V. R.; Lindquist, E.; Payyavula, R. S.; Barros-Rios, J.; Dixon, R.; Engle, N.; Sykes, R. W.; Davis, M.; Jawdy, S. S.; Gunter, L. E.; Thompson, O.; DiFazio, S. P.; Evans, L. M.; Winkeler, K.; Collins, C.; Schmutz, J.; Guo, H.; Kalluri, U.; Rodriguez, M.; Feng, K.; Chen, J.-G.; Tuskan, G. A. A 5-Enolpyruvylshikimate 3-Phosphate Synthase Functions as a Transcriptional Repressor in Populus. *Plant Cell* **2018**, *30*, 1645−1660.

(15) Li, X. Infiltration of Nicotiana benthamiana Protocol for Transient Expression via Agrobacterium. *Bio-protocol* **2011**, *1*, e95.