

Programmed sequential cutting endows Cas9 versatile base substitution capability in plants

Wei Yang^{1†}, Wei Qi^{2†}, Yucai Li^{2†}, Jiyao Wang², Yanmin Luo², Dehui Ding², Sudong Mo²,
Bo Chen², Yu Lu¹, Huarong Li^{2*} & Linjian Jiang^{1*}

¹Key Laboratory of Pest Monitoring and Green Management, Ministry of Agriculture and Rural Affairs; Department of Plant Pathology, College of Plant Protection, China Agricultural University, Beijing 100193, China;
²Qingdao Kingagroot Compounds Co. Ltd., Qingdao 266426, China

Received July 30, 2020; accepted August 7, 2020; published online September 21, 2020

Citation: Yang, W., Qi, W., Li, Y., Wang, J., Luo, Y., Ding, D., Mo, S., Chen, B., Lu, Y., Li, H., et al. (2021). Programmed sequential cutting endows Cas9 versatile base substitution capability in plants. *Sci China Life Sci* 64, 1025–1028. <https://doi.org/10.1007/s11427-020-1798-4>

Dear Editors,

CRISPR/Cas9-mediated genome editing techniques have triggered a revolution in biology research (Jinek et al., 2012). Cas9 coupled with guide RNAs cuts DNA at precise positions, and the resulting double stranded breaks (DSB) were effectively repaired by non-homologous end joining (NHEJ) pathway in higher eukaryotic cells, including animals and plants. As the repair process of NHEJ is error-prone, the resultant outcomes were mostly out-of-frame events. Therefore, Cas9 was primarily considered as a highly-efficient “knock-out” tool, and deeply believed unable to perform in-frame base conversions without significant modifications. As a result, in-frame base changes have to rely on deaminase-mediated base editors (Komor et al., 2016; Gaudelli et al., 2017), prime editing tools (Anzalone et al., 2019), or low-efficient incorporation of donor DNA template via homology directed repair or NHEJ.

Recently, a growing body of evidence revealed that NHEJ repair outcomes are nonrandom and predictable (Shen et al., 2018; Allen et al., 2018; Chen et al., 2019). Indeed, it is well-known that +1/–1 bp indels often dominate the NHEJ repair outcomes even at the same cutting site. We suddenly realized

that a predicted repair outcome could serve as a new DNA target to be edited again mediated by another single guide RNA (sgRNA). Specifically, we hypothesized that when a predicted –1 bp outcome is followed by a +1 bp repair, or vice versa, in-frame base changes could be achieved at a relatively high frequency, and this could become a novel base substitution tool (Figure 1A). Surprisingly, this simple but promising strategy was never explored.

To test this idea, we selected *acetolactate synthase* (ALS) in *Arabidopsis thaliana* as a target gene, where missense point mutations could confer herbicide tolerance. We chose S653 codon AGT to be cut between G and T defined by the downstream NGG PAM (Figure 1B), as S653N (AGT to AAT) mutation confers plant tolerance to imazapic herbicide (Powles and Yu, 2010; in Supporting Information). A pair of sgRNAs were designed and cloned into pHEE401E vector (Wang et al., 2015; in Supporting Information) that should guide Cas9-NGG, driven by a plant egg cell specific promoter, to sequentially cut the site twice in egg cells. Specifically, the NHEJ repair of the DSB mediated by the first sgRNA could delete the G (–G) at the cutting site, and the second sgRNA was designed to target this new resultant sequence, which initiates the second round of NHEJ repair that could add an A (+A) back at the same cutting site (Figure 1B). *Arabidopsis* plants were transformed respectively through floral-dipping with control-653 and CSE-653 vec-

†Contributed equally to this work

*Corresponding authors (Huarong Li, email: lihuarong@kingagroot.com; Linjian Jiang, email: jianglinjian@cau.edu.cn)

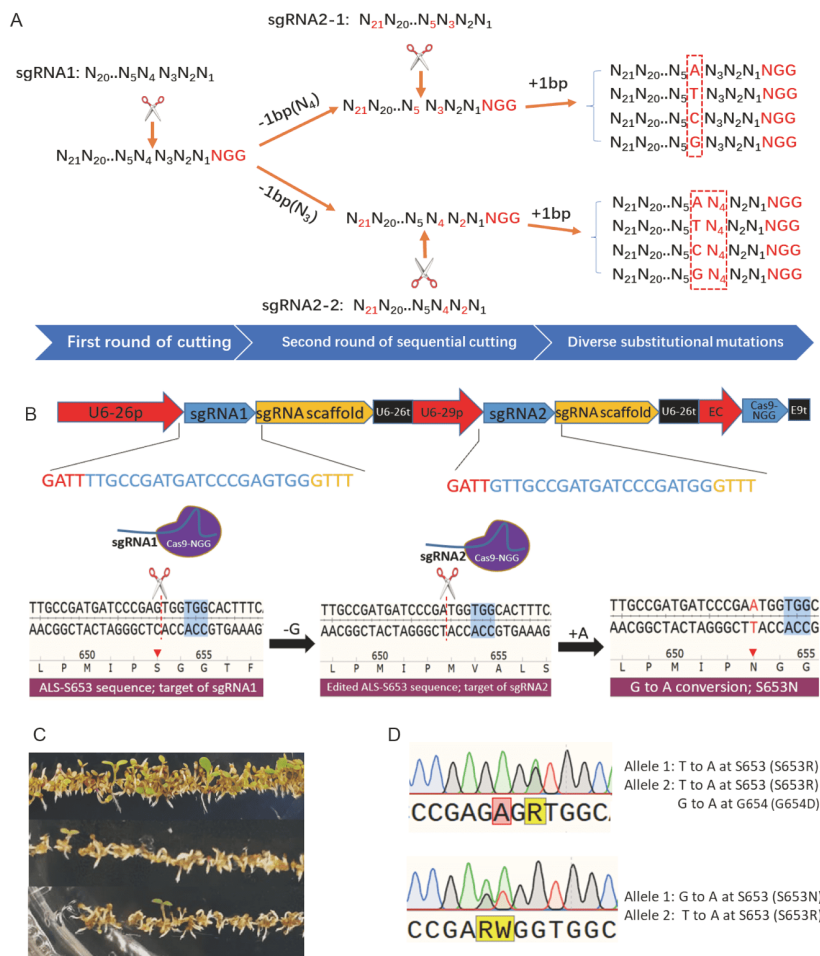


Figure 1 Programmed sequential cutting design endows Cas9 versatile base substitution capability. **A**, Two sequential cuttings by Cas9-NGG could theoretically generate diverse substitutional mutations. NGG in red is the PAM sequence. Potential substitutional mutations are highlighted in red box. Cutting sites are indicated by arrows. **B**, The programmed sequential cutting design aimed to perform G to A base substitution at S653 codon of ALS gene in *Arabidopsis thaliana*. The structure of CSE-653 vector designed to sequentially cut S653 codon twice. pHEE401E backbone was used to ubiquitously express two sgRNAs by U6-26 and U6-29 promoters, and Cas9 was driven by an egg cell specific promoter (EC). PAM is highlighted in blue background. **C**, Herbicide tolerant mutants emerged at T2 generation from three lines upon imazapic selection at 0.24 mg L^{-1} . **D**, Sequencing chromatogram of PCR amplicons from two representative herbicide tolerant mutants. S653N, S653R and S653R/G654D resulted from G to A, T to A conversions. R stands for mixed signals of A and G, and W stands for mixed signals of A and T. Highlighted nucleotides by background color indicate inconsistency with the WT counterparts. Subclones from these PCR amplicons were made to separate each edited allele.

tors. The control-653 vector was designed to produce only the first sgRNA for generation of the –G edit, while the CSE-653 vector to produce both the first and second sgRNAs.

Genotyping showed that the editing efficiency at T1 generation was 66.7% and 60.0% for the control-653 and CSE-653 vectors, respectively (Figure S1A in Supporting Information). All these edited plants had an allele with small indels and a wild type (WT) allele revealed by ICE CRISPR Analysis Tool (Hsia et al., 2018; in Supporting Information). This is reasonable as loss-of-function on both ALS alleles is lethal. Moreover, one plant from control-653 group contained 3% –G edits, suggesting that –G edits did occur in spite of a low frequency. Since re-editing occurs in egg cells (Chen et al., 2017), independent re-editing events could be effectively passed to the next generation, thus T2 seeds were

subject to imazapic selection. No T2 plants from approximately 1,050 seeds of 10 control-653 lines survived. However, seven T2 tolerant mutants out of 1,850 seeds emerged and grew from four out of seven CSE-653 lines (Figure 1C). Interestingly, all these plants contained substitutional mutations, i.e., T to A and G to A conversions, not only on S653 codon but also on G654 codon, which resulted in diverse missense mutations, i.e., S653N, S653R and S653R/G654D (Figure 1D).

Cas9-NG in theory allows three times more editing sites than Cas9-NGG (Nishimasu et al., 2018; in Supporting Information) and thus could endow CSE strategy more flexibility. To examine whether CSE works with Cas9-NG, control-574 and CSE-574 vectors were designed and constructed to cut the ALS-W574 codon twice in *Arabidopsis*

(Figure S2A in Supporting Information). W574L mutation on ALS was known to confer a broader tolerance spectrum to ALS-inhibiting herbicides than S653N, thus is of greater agronomic importance (Powles and Yu, 2010; in Supporting Information). Unfortunately, no suitable base editor was available for such G to T base conversion that leads to W574L mutation. Ten T1 transgenic seedlings from both control-574 and CSE-574 groups were randomly sampled for genotyping, and only two plants from control-574 group were edited but without -G repair (Figure S1A in Supporting Information). The editing efficiency for W574 at T1 generation was lower than for S653, which is not surprising as the efficiency of Cas9-NG at NGA PAM was expected lower than Cas9-NGG (Nishimasu et al., 2018; in Supporting Information). Therefore, T2 seeds were further subject to imazapic selection. No imazapic-tolerant T2 plants emerged from 2,050 seeds of all tested eight control-574 lines; however, from 850 seeds of three tested CSE-574 lines, 71 T2 plants survived imazapic at the frequency of 8.0%–9.2% (Figure S2B in Supporting Information). Sanger sequencing data of 10 randomly sampled herbicide tolerant plants demonstrated that seven plants contained only in-frame substitutional mutations, five of them were homozygous W574L as a result of G to T conversion, one was heterozygous W574L/WT and the remaining one was biallelic W574L/W574M derived from an additional TG to AT conversion (Figure S2C in Supporting Information). The other three plants contained out-of-frame alleles; however, all had in-frame W574L or W574M alleles revealed by sequencing individual amplicon subclones. We speculated that the unusual TG to AT conversion likely resulted from the first round of -G edit, which was then followed by the second round of simultaneous -T/+AT repairs.

Encouraged by the highly efficient base substitution at W574 at T2 generation, we speculated that it could also happen at T1 generation. Therefore, the remaining T1 seeds that would produce about 50 transgenic plants from each of the control-574 and CSE-574 groups were subject to imazapic selection. No control plants survived the herbicide, while four plants did from CSE-574 group (Figure S3 in Supporting Information). Genotyping confirmed that they all harbored W574L or W574M mutations, and two plants had -G edits expected to be cut again mediated by the second sgRNA.

Interestingly, these efficiency data were consistent with the predictions by different online modeling tools including in-Delphi (Shen et al., 2018), FORECast (Allen et al., 2018) and Lindel (Chen et al., 2019). For instance, Lindel predicted that the first round of NHEJ repair at S653 had a 2.3% chance to produce a desired -G edit and the second round had a 7.9% probability to generate the desired +A repair outcome. Thus, it had a 0.18% overall probability for G to A conversion resulting in S653N. In contrast, the corresponding probability to generate the first and second desired edits at W574

was predicted to be 27.8% and 37.9%, respectively, thus 10.5% of all possible repair outcomes would be G to T conversion resulting in W574L. These results suggested that these prediction models, although derived from mammalian cells, likely function in plants, and would be helpful for CSE sgRNA design. Cautiously, unpredicted substitutional mutations, such as TG to AT conversions at W574 codon, suggested that NHEJ repair pattern in plants may need further investigation.

To examine whether CSE works in other plant species, CSE-2038 and control-2038 vectors were constructed targeting the W2038 codon on rice *acetyl-CoA carboxylase 2-like* (ACCCase2) gene (Figure S4A in Supporting Information), as W2038 missense mutations confer herbicide tolerance (Powles and Yu, 2010; in Supporting Information). After transformation, calli were selected on N6 medium supplemented with both hygromycin and quizalofop herbicide. Putatively resistant calli grew from about 1/5 transformed calli with CSE-2038 vector (Figure S4B in Supporting Information), and Sanger sequencing of PCR amplicon revealed a clear base substitution pattern, i.e., 7/15 examined calli had either homozygous or heterozygous W2038L alleles due to programmed G to T conversion (Figures S4C and S5 in Supporting Information). In contrast, less than 1% control-2038 vector transformed calli survived, of which 12 best ones exhibited a typical knockout pattern. Unfortunately, there was an off-target with no mismatches to the spacer, and indeed off-targeted base substitution also occurred (Figure S6 in Supporting Information). Nevertheless, these data indicated that CSE base substitution strategy is applicable to other crops.

Currently, base substitution without a DNA donor is feasible thanks to the invention of base editors (Komor et al., 2016; Gaudelli et al., 2017) and prime editing tools (Anzalone et al., 2019) by addition of new functional domains, such as deaminases and reverse transcriptase, to guide *in vivo* DNA repair processes. However, in countries that regulate gene-edited products based on methods, crops improved using these tools could be subject to strict GMO-like rules. For example, according to the regulatory policy in Australia (Mallapaty, 2019), the resultant base-edited products are regulated as GMOs.

In contrast, by solely relying on the endonuclease activity of Cas9, CSE strategy provides a novel and simpler solution to make versatile base substitution in plants without either donor templates, Cas9-independent off-target (Jin et al., 2019) or guiding repair processes in cells (Table S1 in Supporting Information). Therefore, CSE strategy would be treated as a non-GM method in Australia (Mallapaty, 2019). As this CSE strategy would also function in animal cells, CSE-mediated base substitution could be more attractive to not only crop improvement industry, but also husbandry and aquaculture. Moreover, as other base editors are considerably

larger than Cas9 (Table S1 in Supporting Information), the CSE strategy might find some favorable niches for base substitution in animal cells.

Although the CSE strategy could theoretically convert bases without restrictions, it could suffer low efficiency. However, the CSE strategy allows many different specific ways to achieve base substitution. We demonstrated that base conversions were achieved via deletion of a base followed by insertion of a different one (−1/+1), indicating the opposite way, which is insertion of one base followed by deletion of a different one (+1/−1), should also work. We recommend to utilize the prediction tools, such as Lindel (Chen et al., 2019), to make specific plans using different PAMs. The predictions may suggest that +1/−1 is more promising than −1/+1, or the PAM on the Watson strand is much better than that on the Crick strand. Furthermore, recent breakthrough on PAM expansion (Miller et al., 2020; Walton et al., 2020) could also be compatible with CSE strategy, which could further increase the flexibility of CSE.

In conclusion, our study showed the sequential editing design could turn Cas9 into a versatile base editor in plant cells. As CSE editing could be deployed simply by including an additional synthetic sgRNA, we believe this novel base substitution strategy could be applied to perform flexible base substitution in plants and other organisms.

Compliance and ethics A patent was filed to patent office in China, and key authors were listed as inventors. L.J. and H.L. designed the experiments; W.Y., W.Q. and Y.L. performed most of the experiments with the help from other authors; H.L. supervised the project at Kingagroot and L.J. supervised the project at CAU; H.L. and L.J. wrote the manuscript; L.J. conceived the idea.

Acknowledgements This work was supported by the National Transgenic Science and Technology Program (2019ZX08010-003) and National Natural Science Foundation of China (31872933).

SUPPORTING INFORMATION

The supporting information is available online at <https://doi.org/10.1007/s11427-020-1798-4>. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

- Allen, F., Crepaldi, L., Alsinet, C., Strong, A.J., Kleshchevnikov, V., De Angeli, P., Páleniková, P., Khodak, A., Kiselev, V., Kosicki, M., et al. (2018). Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat Biotechnol* 37, 64–72.
- Anzalone, A.V., Randolph, P.B., Davis, J.R., Sousa, A.A., Koblan, L.W., Levy, J.M., Chen, P.J., Wilson, C., Newby, G.A., Raguram, A., et al. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149–157.
- Chen, W., McKenna, A., Schreiber, J., Haeussler, M., Yin, Y., Agarwal, V., Noble, W.S., and Shendure, J. (2019). Massively parallel profiling and predictive modeling of the outcomes of CRISPR/Cas9-mediated double-strand break repair. *Nucleic Acids Res* 47, 7989–8003.
- Chen, Y., Wang, Z., Ni, H., Xu, Y., Chen, Q., and Jiang, L. (2017). CRISPR/Cas9-mediated base-editing system efficiently generates gain-of-function mutations in *Arabidopsis*. *Sci China Life Sci* 60, 520–523.
- Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* 551, 464–471.
- Jin, S., Zong, Y., Gao, Q., Zhu, Z., Wang, Y., Qin, P., Liang, C., Wang, D., Qiu, J.L., Zhang, F., et al. (2019). Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science* 364, 292–295.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821.
- Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424.
- Mallapaty, S. (2019). Australian gene-editing rules adopt ‘middle ground’. *Nature* <https://doi.org/10.1038/d41586-019-01282-8>.
- Miller, S.M., Wang, T., Randolph, P.B., Arbab, M., Shen, M.W., Huang, T. P., Matuszek, Z., Newby, G.A., Rees, H.A., and Liu, D.R. (2020). Continuous evolution of SpCas9 variants compatible with non-G PAMs. *Nat Biotechnol* 38, 471–481.
- Shen, M.W., Arbab, M., Hsu, J.Y., Worstell, D., Culbertson, S.J., Krabbe, O., Cassa, C.A., Liu, D.R., Gifford, D.K., and Sherwood, R.I. (2018). Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* 563, 646–651.
- Walton, R.T., Christie, K.A., Whittaker, M.N., and Kleinstiver, B.P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* 368, 290–296.