COMMENTARY

Programmable DNA transposon-mediated targeted gene insertion in plants

Yingbo Liang[1,2] **and Yanpeng Wang**[1,2,3,*]

[1]Key Laboratory of Seed Innovation, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China
[2]Center for Genome Editing, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China
[3]College of Advanced Agricultural Sciences, University of Chinese Academy of Sciences, Beijing 100101, China
*Correspondence: yanpengwang@genetics.ac.cn

Received: July 16, 2024; Accepted: August 26, 2024; Published Online: October 16, 2024; https://doi.org/10.59717/j.xinn-life.2024.100095

Citation: Liang Y. and Wang Y. (2024). Programmable DNA transposon-mediated targeted gene insertion in plants. The Innovation Life **2(4)**: 100095.

Since the pioneering genetic transformation of plants using *Agrobacterium* was first reported in 1983, transgenic crops have been cultivated in over 40 countries, significantly contributing to global crop production to meet the demands of a growing population. However, conventional plant transgene integration methods, such as *Agrobacterium*-mediated and particle bombardment-based transformation, typically integrate foreign genetic elements at random locations. This randomness can sometimes lead to undesired mutations that affect plant traits. So, targeted gene insertions have been highly desired and actively pursued in plants due to their significant potential for plant trait improvement and synthetic biology systems.[1]

Recently, the study by Liu et al. describes a new genome engineering tool that can efficiently perform targeted gene insertion in plant genome. Taking advantage of the ability of transposable elements (TEs) to excise and integrate DNA into the genome, the authors combined rice *Pong* DNA transposon system and Cas protein, thereby generating a transposase-assisted target-site integration (TATSI) system (Figure 1), enabling programmable transposase-mediated targeted DNA insertion in *Arabidopsis* and soybean with high efficiency and precision.[2]

Over the past ten years, various genome editing-based approaches have been utilized for targeted gene insertions in plants. One notable approach is

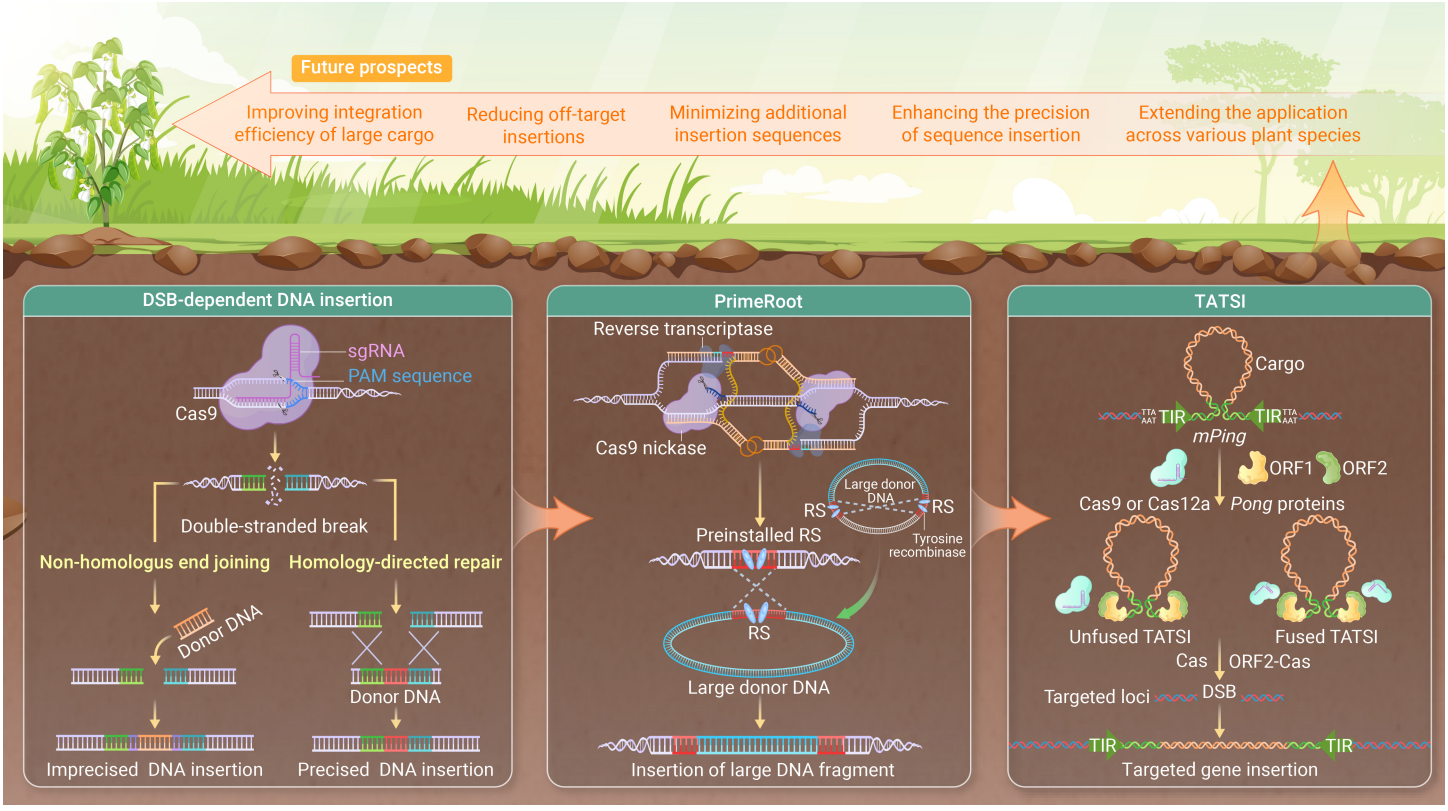

Figure 1. Overview of targeted DNA insertion in plants Schematic of programmable DNA insertion with NHEJ, HDR, PrimeRoot and TATSI. When DSBs are produced by sequence-specific nucleases (SSNs) at the specific target site, the presence of donor DNA facilitates the repair of DNA break, resulting in the insertion of a targeted DNA sequence. In NHEJ, the repair is error-prone, often accompanied by the insertion or deletion of bases at the junction. In HDR, the repair of double-strand DNA break is facilitated by homologous donor DNA template. HDR can result in precise DNA insertions, but the efficiency of HDR in plant cells is extremely low. The PrimeRoot system is composed of an enhanced plant prime editor, an optimized prime editing guide RNA (pegRNA), and an enhanced tyrosine recombinase. The dual-ePPE inserts the recombination site (RS) sequence in target site. Meanwhile, the recombinase identifies and excises two identical RS sites on the donor vector, resulting in an intermediate donor containing the donor DNA fragment with one corresponding RS. Fallowing this, the donor DNA fragment is recombined into the inserted RS, generating a precise large DNA insertion at the target site. In TATSI system, cargos were embedded into *mPing* element, generating synthetic *mPing*_cargo element. ORF1 and ORF2 proteins from the *Pong* transposon were combined with nucleases (Cas9 or Cas12a), which generated TATSI system. *mPing*_cargo element was excised by transposase precisely and target site was cleaved by nucleases, finally, *mPing*_cargo integrated at the target site. Future developments aim to enhance the efficiency and safety of genetic modifications, ultimately contributing to the creation of crops with improved traits such as higher yield, better disease resistance, and enhanced nutritional content.

depended on the nuclease-induced DNA double-strand breaks (DSBs), which can induce gene insertion via non-homologous end joining (NHEJ) or homology-directed repair (HDR) when a donor DNA template is provided. NHEJ pathway is highly imprecise, often resulting in random DNA base insertions and deletions (indels) at the ends of the inserted donor fragment. In HDR, several strategies have been employed to enhance its efficiency in plants, such as the use of positive-negative selection or increasing the amount of donor template through viral vectors. Despite these efforts, HDR-mediated

gene insertion remains quite challenging in plants.[3] The recently developed PrimeRoot (prime editing-mediated recombination of opportune targets) system using prime editing (PE) to direct site-specific recombinases (SSRs) for programmable large DNA insertions in plants. SSRs, such as Cre and FLP, have been widely used for genome manipulation in the past, relying on their specific recombination sites. PrimeRoot system can first install recombination sites at the desired target location using an optimized prime editor. Then, it achieves precise integration of large DNA sequences when donor DNA and recombinase are provided. This system has successfully integrated large DNA fragments of up to 11.1 kb in rice genome and remarks the beginning of precise large DNA manipulation in plants (Figure 1).[4]

Transposons are naturally mobile DNA sequences capable of independently transferring large DNA cargoes within the host genome. Many transposons have been developed for large DNA insertion in bacteria and mammalian cells, such as CRISPR-associated transposases (CASTs) and R2 retrotransposons. More specifically class II DNA transposons excise their corresponding TEs DNA from one site and then precisely integrate the TEs DNA into a new site within the genome. *mPing* is an active miniature inverted-repeat transposable element (MITE) firstly found in the genome of rice (*Oryza sativa*). It is approximately 430 base pairs in length, featuring 15 bp terminal inverted repeats (TIRs) and specific target site duplications (TSDs), but it lacks open reading frames. The transposition of *mPing* is associated with the activity of the *Pong* transposase. *Pong* is a longer transposon, measuring 5,166 base pairs, and shares similar TIRs and subterminal regions with *mPing*. *Pong* contains two open reading frames (ORFs). ORF1 has weak sequence similarity to the DNA-binding domains of Myb and attaches to the 15 bp TIRs at the ends of the *mPing* element. ORF2 includes a putative Asp-Asp-Glu (DDE) motif, which is part of the catalytic active center of transposases found in many bacterial insertion sequences and some eukaryotic transposons. The non-autonomous transposable element *mPing*, flanked by TTA/TAA repeats, requires *Pong*'s ORF1 and ORF2 for its excision and integration processes and shows a preference for inserting into TAA/TTA sequences in the genome.[5]

Building on the "cut-and-paste" characteristic of the *mPing/Pong* transposon system, Liu et al. fused the rice *Pong* transposase protein to CRISPR nucleases to develop a programmable TATSI system (Figure 1). The excision of *mPing* within the GFP expression cassette in the *Arabidopsis* genome confirmed that *mPing* could be excised when ORF1 and ORF2 were present, regardless of Cas9 fusion, indicating the transposition capability of ORF1 and ORF2. Targeted insertion of *mPing* at endogenous gRNA target sites confirmed that targeted insertion occurring only in the presence of active Cas9 and the ORF1+ORF2 proteins. Additionally, LbCas12a was also shown to be effective in combination with ORF1+ORF2 proteins for targeted integration (Figure 1).[2]

In *Arabidopsis*, by altering the targeting sequence of the CRISPR gRNA, TATSI efficiently achieved targeted insertion of the *mPing* transposon at various gene loci. In soybean, by adjusting the design of the gRNA and the transposon vector, TATSI also achieved efficient targeted integration. Additionally, TATSI facilitated the targeted insertion of enhancer elements, open reading frames, and gene expression cassettes in *Arabidopsis* and soybean by embedding these sequences into the *mPing* element, and the size of cargo is up to 8.6 kb. In addition, the TATSI system not only increased the frequency and accuracy of targeted integration of long sequences but also reduced the preference for insertions of the *mPing/Pong* transposon system at TAA/TTA sites.[2] The successful application of TATSI technology in the model plant *Arabidopsis* and the staple food crop soybean demonstrates its broad application prospects in crop improvement. With continuous optimization and expansion of its application scope, TATSI is expected to play a crucial role in upgrading molecular breeding technologies and benefitting future agriculture.

Despite significant progress having been made with TATSI, there remain some technical issues to be resolved. For instance, TATSI currently relies on DSBs for targeted DNA insertion, and there is a need to further improve accuracy and reduce off-target insertions resulting from random transpositions. Additionally, optimizing the *mPing* transposon vector to carry longer DNA sequence and controlling the orientation of *mPing* insertion while minimizing the additional undesired insertion sequences of *mPing* itself is crucial. Extending the application of TATSI across a variety of crop species also poses a significant challenge. Future research is expected to intensively explore these aspects to enable the widespread application of TATSI technology.

REFERENCES

1. Dong, O.X. and Ronald, P.C. (2021). Targeted DNA insertion in plants. Proc. Natl. Acad. Sci. USA **118**: e2004834117. DOI: 10.1073/pnas.2004834117.
2. Liu, P., Panda, K., Edwards, S.A., et al. (2024). Transposase-assisted target-site integration for efficient plant genome engineering. Nature **631**: 593–600. DOI: 10.1038/s41586-024-07613-8.
3. Gao, C. (2021). Genome engineering for crop improvement and future agriculture. Cell **184**: 1621–1635. DOI: 10.1016/j.cell.2021.01.005.
4. Sun, C., Lei, Y., Li, B., et al. (2024). Precise integration of large DNA sequences in plant genomes using PrimeRoot editors. Nat. Biotechnol. **42**: 316–327. DOI: 10.1038/s41587-023-01769-w.
5. Jiang, N., Bao, Z., Zhang, X., et al., (2003). An active DNA transposon family in rice. Nature **421**: 163-167. DOI: 10.1038/nature01214.

FUNDING AND ACKNOWLEDGMENTS

This work was supported by the Biological Breeding-Major Projects (2023ZD0407403), the National Natural Science Foundation of China (32122051), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA0450301), the Youth Innovation Promotion Association of the Chinese Academy of Sciences (2020000003). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

Y.L. and Y.W. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DATA AND CODE AVAILABILITY

Not applicable.