Molecular engineering

https://doi.org/10.1038/s41477-023-01496-9

The power of repetition

Holger Puchta

The inclusion of retrotransposon long terminal repeats — and of other repeated sequences — enhances transfer DNA copy numbers in plant cells during transformation. Gene editing and homologous recombination-mediated gene targeting can therefore be improved by these means: however, the mechanism remains a mystery.

The application of CRISPR–Cas-based genome-engineering technologies has revolutionized plant biology and breeding. In this issue of *Nature Plants*, Dickinson et al.[1] report on a strategy for how to improve the efficiency of these approaches.

The most popular way of transforming plants is using the soil bacterium *Agrobacterium tumefaciens*, which transfers a specific piece of single-stranded DNA (the transfer DNA (T-DNA)) from its tumour-inducing (Ti) plasmid into plants cells. Sequence-specific borders define which sequences are transferred to the plant cell and later integrated into the plant genome[2]. By simple insertion between these borders, any gene of interest can be transferred to plants. Thus, gene editing tools (such as a Cas nuclease and its respective single-guide RNA) can be transferred. The T-DNA is covalently linked via a tyrosine bond to the bacterial VirdD2 protein, which targets the T-stand (the single-stranded form of T-DNA) to the plant nucleus. In the nucleus, the T-DNA becomes double-stranded either before or during the integration. Double-stranded T-DNA is transcribed even if integration does not occur. Recent results have shown that integration is mediated by the DNA repair factors of the host[3]. The free 3′ end of the T-DNA is linked to a broken chromosomal DNA end by polymerase-θ (POLQ) using microhomologies. The VirD2 protein is either removed by hydrolysis or the 5′ end is processed by the DNA nuclease MRE11 before it is linked to the other broken chromosomal end by non-homologous end joining[3].

T-DNA integration occurs randomly in the genome. However, it has previously been shown that homology-directed integration at a specific site can be massively enhanced by the induction of a double-strand break[4]. Nevertheless, in comparison to random integration, the frequencies obtained are, at most, in the low per cent range and various efforts have been undertaken over the past 25 years to improve gene-targeting technology in plants[5]. The group of D. Voytas were able to show that by using the replicase of a *Geminivirus* to amplify the copy number of the template, gene targeting can be massively enhanced in plants cells[6].

Dickinson et al. speculated that by replacing a part of a functional retrotransposon with a repair matrix, it might also be possible to increase template numbers. Indeed, a T-DNA that carries the two long terminal repeats (LTRs) of the active ONSEN element showed a strong increase in copy numbers (up to 50-fold) in plant cells as measured by quantitative PCR, a result that was verified by fluorescence in situ hybridization analysis. The big surprise came when the authors tried to define the sequence-specific requirements for this amplification in more detail. As expected, replacing the LTRs of ONSEN by those of other retrotransposons led to similar amplification numbers. If one or the

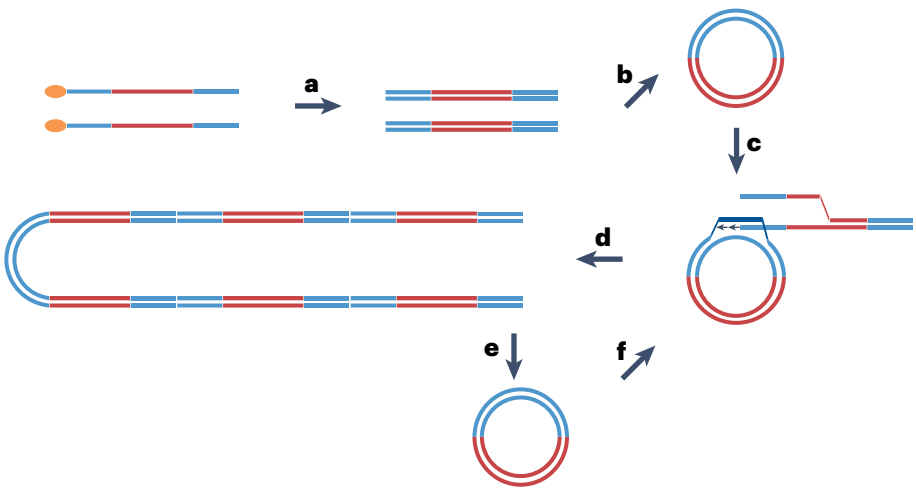

Fig. 1 | An alternative model for repeat-induced DNA amplification. a, T-DNAs, as single-stranded molecules attached to the VirD2 protein (in orange), are transferred to the plant cell nucleus where they become double-stranded. The VirD2 protein is removed by the action of MRE11. **b**, Intramolecular ligation of the free ends of a T-DNA by POLQ leads to formation of a circle. **c**, By invading this double-stranded circle with the 3′ end of a repeat of another T-DNA, a D-loop is formed that might activate the cellular response for replicative DNA damage (which involves ATR and RAD17). **d**, The induced response leads to replicative repair, which results in a rolling circle-based amplification process of tandem repeats of the respective T-DNA. **e**,**f**, Further circles might arise from tandems repeats (**e**), which then might lead to further amplification (**f**). Repeats are depicted in blue; the unique sequence is depicted in red.

other single LTR was removed, the increase in copy numbers was much lower. However, the envisaged 'negative control' (repeats of similar size with no relation to retrotransposons) also induced amplification with high efficiency. This applied to different kinds of repeats with different lengths, as well as GC contents.

For testing whether gene targeting can indeed be improved in this way, the authors used the well-established *in planta* gene-targeting assay in *Arabidopsis*. Here, the template is first integrated into the genome and double-strand breaks are not only induced in the target locus but also in the transgene to cut out and activate a repair matrix during plant development. An organ-specific promoter can be used to restrict the reaction to the egg cell, which guarantees heritability of the respective change[7]. Including the ONSEN LTRs at both ends of the cut-out fragment enhanced the gene-targeting frequency by threefold. It is possible that this enhancement is also due to repeat-induced copy number amplification of the chromosomal template after excision. This indicates that the phenomenon of repeat-induced amplification might not be restricted to T-DNAs.

This raises the questions of how these intriguing results can be explained, and what mechanism lies behind the amplification. The authors were able to shed some light on the process. They showed that – at least parts of – the amplified T-DNAs are concatenated and that microhomologies are involved in their formation, which points to a role of POLQ in the reaction. However, the nature of T-DNA junctions does not allow conclusions about whether the amplified DNAs are chromosomal or extrachromosomal and whether they are linear arrays or circles of one or several monomers. The authors further identified three factors that are required for amplification: RAD17 and the kinase ATR (both of which are involved in the cellular response to replication damage), as well as the nuclease MRE11 (which is required for VirD2 removal). It is tempting to speculate that interactions between repeats might enable the build-up of structures that mimic replicative DNA damage, which are then dissolved by the cellular machinery via an amplification step. The authors suggest a model in which concatenations occur during repeated cycles of T-DNA integration at the same site. Alternatively, it is worth considering that an active amplification step occurs in plant cells extrachromosomally before integration (Fig. 1). A very efficient form of amplification is amplification by rolling circle replication, which has been described for plasmid and viral replication (including geminiviruses): a prerequisite for this is the formation of a circular intermediate. Indeed, T-DNA circles are formed in plant cells during transformation[8], and the circularization of linear DNA might also be regularly associated with double strand break-induced fragment excision as in *in planta* gene targeting.

Dickinson et al. raise more questions than they answer, but the implications of their study for practical applications are manifold. The article will definitely inspire many scientists to perform follow-up experiments to answer the remaining questions, including whether the phenomenon is reproducible in other plant species besides *Arabidopsis*; how small the repeats can be to still induce amplification; whether we can use this approach to boost the transient expression of proteins in general; and whether this phenomenon occur naturally, as circular extrachromosomal DNA-carrying repeated sequences are regularly found – not only in plants but also other in eukaryotes[9].

Holger Puchta
Joseph Gottlieb Kölreuter Institute for Plant Sciences (JKIP), Karlsruhe Institute of Technology, Karlsruhe, Germany.
e-mail: holger.puchta@kit.edu

Published online: 31 August 2023

References

1. Dickinson, L. et al. *Nat. Plants* https://doi.org/10.1038/s41477-023-01495-w (2023).
2. Gelvin, S. B. *Annu. Rev. Genet.* **51**, 195–217 (2017).
3. Kralemann, L. E. M. et al. *Nat. Plants* **8**, 526–534 (2022).
4. Puchta, H., Dujon, B. & Hohn, B. *Proc. Natl Acad. Sci. USA* **93**, 5055–5060 (1996).
5. Dong, O. X. & Ronald, P. C. *Proc. Natl Acad. Sci. USA* **118**, e2004834117 (2021).
6. Baltes, N. J., Gil-Humanes, J., Cermak, T., Atkins, P. A. & Voytas, D. F. *Plant Cell* **26**, 151–163 (2014).
7. Wolter, F., Klemm, J. & Puchta, H. *Plant J* **94**, 735–746 (2018).
8. Bakkeren, G., Koukolíková-Nicola, Z., Grimsley, N. & Hohn, B. *Cell* **57**, 847–857 (1989).
9. Peng, H., Mirouze, M. & Bucher, E. *Curr. Opin. Plant Biol.* **69**, 102263 (2022).

Competing interests

The author declares no competing interests.