scientific reports

Check for updates

OPEN

Double step screening using endogenous marker improves relative gene targeting efficiency in Arabidopsis

Yiqiu Cheng[1,2], Lei Zhang[1,2], Yongping Ke[1,2], Xiaofei Dang[1] & Daisuke Miki[1,3]✉

Gene targeting (GT) is a powerful tool for manipulating endogenous genomic sequences as intended. However, its efficiency is rather low, especially in seed plants. Numerous attempts have been made to improve the efficiency of GT via the CRISPR/Cas systems in plants, but these have not been sufficiently effective to be used routinely by everyone. Here, we report a surrogate screening method that improves the relative efficiency of CRISPR/Cas9-mediated GT in Arabidopsis. Our findings indicate that simultaneous mutagenesis of the endogenous *MAR1* gene, which results in kanamycin resistance, can be employed to efficiently screen for precise and heritable GT events at multiple endogenous sites in the Arabidopsis genome. In this study, we demonstrate that a double-step screening strategy can achieve up to a four-fold increase in the efficiency of GT in Arabidopsis. The principle of this surrogate system has the potential to be widely applied.

Keywords Genome engineering, Gene targeting, CRISPR/Cas9, *MAR1*, Co-editing, Surrogate system, *Arabidopsis thaliana*

Precise gene targeting (GT), such as knock-in (KI) and sequence replacement, represents a robust tool in genome engineering for molecular biology research and molecular breeding. Nevertheless, GT remains a significant challenge, particularly in seed plants, due to the low efficiency of homologous recombination (HR)[1]. Recently, the use of sequence-specific nucleases (SSNs) to create target site-specific double-strand breaks (DSBs) has been reported to enhance the efficiency of HR[2]. DSBs generated by SSNs are predominantly repaired by error-prone non-homologous end-joining (NHEJ), but rarely by error-free homology-directed repair (HDR) if an appropriate donor repair template is provided[3–6]. Consequently, HDR-mediated GT using SSNs has been documented in numerous plant species, including *Arabidopsis thaliana* (Arabidopsis) and rice[4,7–15]. Nevertheless, the efficacy of precise GT is not yet sufficient for widespread application.

A multitude of endeavors have been undertaken with the objective of enhancing the efficacy of GT in plants. These studies indicate that the frequency of DSBs induced by clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) (hereafter Cas9) is a crucial factor in the successful establishment of GTs[10,16]. Therefore, the most straightforward and effective method for improving GT is to enhance the DSB efficiency of SSNs[14,15]. A sequential transformation strategy has been reported that can yield precise Cas9-mediated GT events in rice and Arabidopsis with high efficiency[7,9]. In brief, in Arabidopsis, donor constructs with single guide RNA (sgRNA) expression cassettes are transformed into parental lines that stably express Cas9 in egg cells and early embryos by the DD45 promoter. The sequential transformation strategy also serves to enhance the frequency of DSBs. This is because the use of highly efficient parental lines allows Cas9 to maintain higher levels of DSB activity. Similar endeavors to enhance GT efficacy at elevated DSB frequencies encompass the utilization of other CRISPR/Cas systems, including Cas12a, and transcription and translation enhancers[8,11,14,17–19]. Nevertheless, the observed improvement in DSB frequency appears to be approaching a plateau, suggesting that further enhancement may be challenging to achieve. Consequently, it is of the utmost importance to develop effective strategies to enhance GT efficiency.

An alternative approach to enhancing the relative efficiency of GT in plants would be to improve the screening performance. One potential avenue for enhancing the relative efficiency of GT would be the implementation of

[1]Shanghai Center for Plant Stress Biology, CAS Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai 200032, China. [2]University of Chinese Academy of Sciences, Beijing 100049, China. [3]Kihara Institute for Biological Research, Yokohama City University, Yokohama, Kanagawa 244-0813, Japan. ✉email: daisukemiki@cemps.ac.cn

a surrogate screening system. Surrogate reporter systems can be employed to enrich cells and plants that have been mutated at target sites by SSNs. Consequently, the simultaneous editing of an endogenous marker gene and subsequent selection according to the mutant phenotype may potentially enhance the efficiency of screening. In other words, if mutant phenotypes generated by genome editing can be employed for screening, plants with high DSB activity would be enriched. While various surrogate reporter systems have been developed to enrich cells with high DSB activity in mammals[20,21], few have been developed in plants[13,22–24]. Moreover, surrogate systems have been scarcely employed in the establishment of GT events in rice[13].

The endogenous gene *Multi-Antibiotic Resistance 1* (*MAR1*)/*RNA-mediated transcriptional gene silencing 3* (*RTS3*) encodes a putative chloroplast and mitochondria transporter in Arabidopsis. The original *RTS3*, which has since been renamed *MAR1*, was initially identified through forward genetic screening, with the objective of finding mutants that reactivate transcriptionally silenced kanamycin resistance *NPTII* transgenes. However, further research yielded unexpected results, indicating that the loss of function of *MAR1* alone is sufficient to confer kanamycin resistance in Arabidopsis[25]. It seems reasonable to conclude, therefore, that *MAR1* is likely to prove a useful marker gene for surrogate systems, given that the mutant displays a kanamycin-resistant phenotype. It is anticipated that when the target site in the *MAR1* gene is cleaved by Cas9, biallelic and fully chimeric mutants will exhibit a kanamycin-resistant phenotype, which allows for the quantification of DSB activity.

The *MAR1*-mediated surrogate approach was initially employed to enhance the editing frequency in genes of interest in T2 Arabidopsis and T0 tomato generations[23]. This study revealed that the Cas9-mediated mutation efficiency was 2.5 to 3 times higher in Arabidopsis T2 generation. Furthermore, all seven regenerated shoots from kanamycin-resistant callus exhibited biallelic and chimeric mutation patterns in T0 tomato. Their findings indicated a correlation between Cas9-mediated mutations in two distinct loci that are targeted simultaneously in plants. Consequently, it is proposed that co-editing can be employed as an approach to enrich plants containing a desired target gene editing.

The present study aimed to develop a surrogate screening system for efficient screening of GT in plants. This was achieved by mutagenesis using the endogenous *MAR1* gene as a reporter to evaluate the DSB activity of Cas9. However, the direct screening of the *MAR1* gene mutation phenotype in the T1 generation Arabidopsis plants did not yield the desired results in the present study. Consequently, a double-step screening process was implemented, whereby transgenic plants were initially selected for Basta resistance, followed by kanamycin selection for the *MAR1* mutation. This approach was employed with the objective of enhancing the efficiency of GT screening. The double-step surrogate system demonstrated an enhanced relative GT efficiency of up to 4.3-fold. However, it should be noted that the survival of some precise GT plants was unsuccessful during the surrogate screening process.

Results

Simultaneous co-editing at two distinct loci

In order to establish an efficient surrogate screening system, a statistical analysis was performed to examine the relationship between mutation rates at two different loci targeted simultaneously by Cas9 (Fig. 1). Two distinct loci were targeted for co-editing. The two loci under consideration were *Salt Overly Sensitive 1* (*SOS1*) and *MAR1* (Fig. 1a). The *SOS1* locus has been reported to yield precise GT with high efficiency[10]. The *SOS1* and *MAR1* co-editing construct, comprising two sgRNA expression cassettes targeting each gene, was transformed into the DD45 promoter::Cas9 parental line (see Methods) via Agrobacterium, in accordance with the established sequential transformation strategy[7]. The transgenic plants were initially screened for Basta resistance via the T-DNA and then selected based on the kanamycin-resistant phenotype of the endogenous *MAR1* gene mutation (Fig. 1b). It is postulated that this double-step screening will facilitate the enrichment of plants exhibiting heightened DSB activity. The mutation rates at these target loci were determined by TIDE[26] (see Methods) for each of the 40 randomly selected plants from approximately 100 independent Basta-resistant T1 transformants. The results of the Cas9-mediated simultaneous editing revealed a positive correlation between the mutation rates at distinct *SOS1* and *MAR1* loci (Fig. 2a). Subsequently, all Basta-resistant T1 transformants were subjected to kanamycin selection via spraying (Fig. 1b). A total of 32 plants were found to be resistant to kanamycin. Twenty-three of the plants exhibited a mutation frequency of nearly 100% at the *MAR1* locus, while others displayed a mutation frequency of 50% or less (Fig. 2b). Moreover, biallelic and heterozygous *SOS1* mutants appeared to be enriched after kanamycin selection (Fig. 2b).

To ascertain the generality of the relationship between the mutation rates of the two target loci, three additional gene regions, *Target of Rapamycin* (*TOR*), and the N- and C-terminal ends of *Demeter* (*DME*), were selected as target sites of interest. Three constructs were created, each with two sgRNAs targeting one of the target regions of interest and *MAR1*. These constructs were transformed into the DD45 pro::Cas9 parental line (Fig. 1a). Two biological replicates were conducted on the three aforementioned constructs, resulting in nearly 100 independent Basta-resistant T1 transformants in each experiment. The mutation rates at the target loci were determined for 20 randomly selected Basta-resistant plants in each experiment. A weak positive correlation was observed between the target sites (the N- and C-terminal ends of *DME*) and *MAR1*, whereas a low mutation frequency was noted at the *TOR* locus (Fig. 2c,e,g). Subsequently, all Basta-resistant transgenic T1 plants were subjected to kanamycin selection, and genotyping analysis was performed on 15 randomly selected kanamycin-resistant plants. Following kanamycin selection, the majority of plants exhibited near-complete mutations at the *MAR1* locus, accompanied by an elevated mutation rate for the gene of interest (Fig. 2d,f,h). In particular, mutant plants targeting the N-terminal region of *DME* were clearly enriched following kanamycin selection (Fig. 2f).

Given that the DD45 promoter is expressed in a single cell stage, specifically in the egg cell and early embryo, the majority of mutations obtained would be heterozygous (mutation rate: 0.5) or biallelic (mutation rate: 1.0),

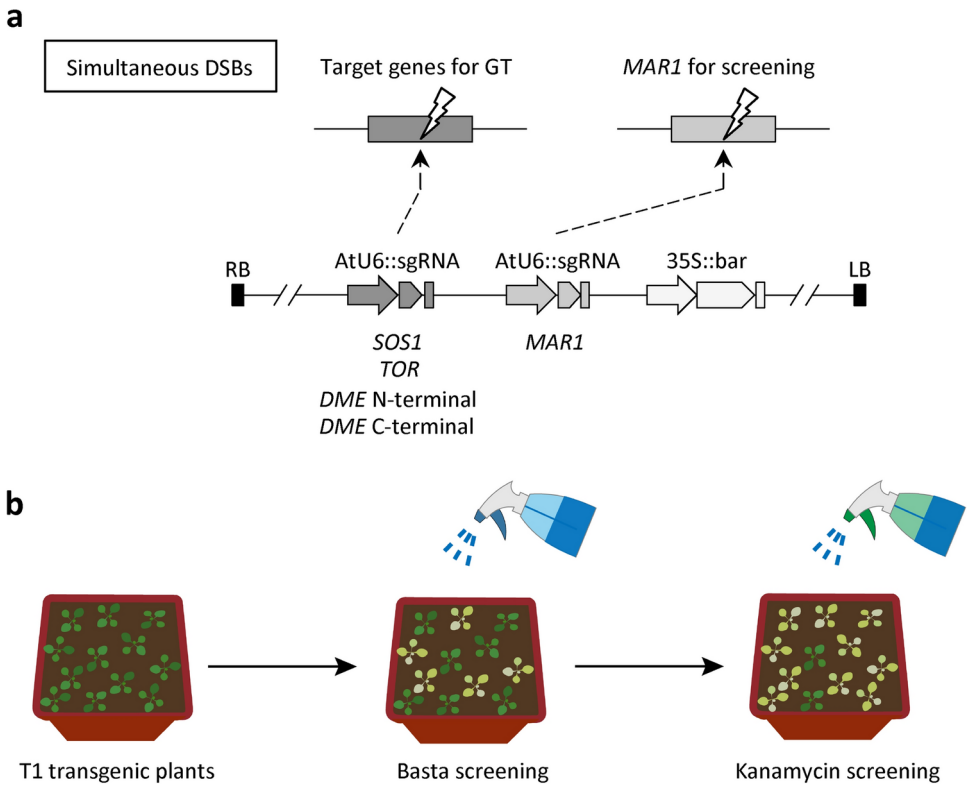

Fig. 1. Simultaneous co-editing surrogate screening system. (**a**) A schematic representation of the simultaneous co-editing process for the purpose of efficient screening. The two sgRNAs in the construct target the gene of interest and the *MAR1* locus. In each construct, the two sgRNAs can simultaneously generate DSBs at two distinct loci. (**b**) Overview of the co-editing surrogate double-step screening system. T1 transgenic plants were germinated directly on soil and subsequently screened by spraying with Basta and kanamycin.

with a few possibilities of chimeric mutations (Fig. 2a–h). Statistical analysis revealed a significant enrichment of mutations in target genes after kanamycin screening, with the exception of the *DME* C-terminal region (Fig. 2i). These results strongly indicate that the double-step surrogate screening system can enrich plants with high DSB activity at a wide range of target sites. The efficiency of DSBs by the CRISPR/Cas systems has been demonstrated to be a crucial factor in establishing precise GT events[10]. Consequently, it is postulated that the relative efficiency of GT in plants can be enhanced by concurrent DSB at the target locus of GT and *MAR1*, followed by kanamycin screening.

Surrogate screening improves GT efficiency

To test this hypothesis, the *GFP* knock-in (KI) donor constructs of *SOS1* flanked by 1 Kbp homology arms were cloned into the mutagenesis vector depicted in Fig. 1a. Constructs containing sense (*SOS1-GFP-S*) or antisense (*SOS1-GFP-AS*) *SOS1* sgRNA recognition sites at both ends of the donor sequence for donor excision were prepared (Fig. 3a, Supplementary Fig. S1a). It is anticipated that the excision of the donor template will result in an enhanced GT efficiency[4,27,28]. Two biological replication experiments were conducted for each construct. Approximately 100 to 200 independent Basta-resistant T1 transgenic plants were obtained in each experiment. To ascertain the efficacy of GT, genotyping analysis was conducted on all the Basta-resistant plants obtained. Full-length primers were designed to anneal upstream and downstream of the homology arms, enabling the amplification of both endogenous and precise GT alleles (Supplementary Fig. S1a). GT events detected with the full-length primer set have been reported to be precisely incorporated by HDR at both homology arms and inherited by the offspring[9–11,16]. At least one plant exhibiting precise *SOS1-GFP* GT positivity was identified, with an efficiency of 1–5% (Table 1). Following kanamycin screening, approximately three-fourths of the T1 plants were killed (Fig. 3b), resulting in an increase in relative GT efficiency of 4–15% (Table 1). It was observed that some of the precise *SOS1-GFP* GT-positive plants did not survive the kanamycin treatment (Table 1). Genotyping results indicated that kanamycin-sensitive plants exhibited a lower mutation frequency at the *MAR1* locus, despite exhibiting precise GT positivity (Fig. 3c). Furthermore, in Experiment #4, three heterozygous *mar1* mutants (one of which was biallelic precise *GFP*-KI) were selected with kanamycin (Fig. 3c).

In addition to *SOS1-GFP*, the *Flag* epitope tag KI to *Repressor of Silencing 1* (*ROS1*) (*ROS1-Flag*) and to the N-terminus of *DME* (*Flag-DME*) and the luciferase KI to the *NITRATE REDUCTASE 2* (*NIA2*) locus (*NIA2-Luc*) were also examined (Fig. 3a). In all experiments, genotyping of Basta-resistant T1 transgenic plants yielded precise KI GT plants with an efficiency of 0.5–3.7% (Fig. 3c, Table 1). As with the *SOS1-GFP* strains, approximately three-quarters of the Basta-resistant T1 strains were unable to survive kanamycin selection. Consequently, the

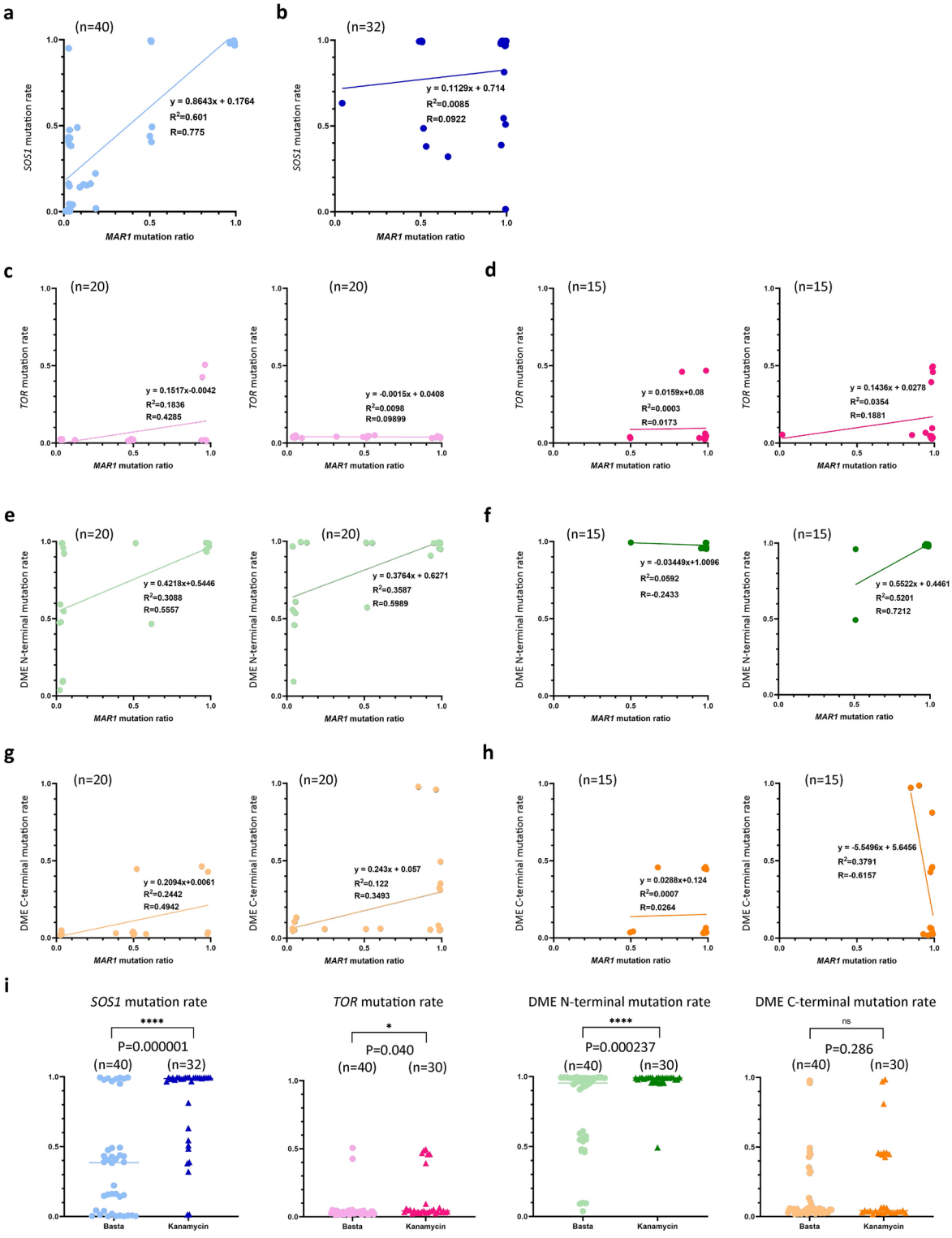

Fig. 2. Statistical analysis of mutation rates by double-step screening. The mutation rates at endogenous target loci were analyzed. The mutation rates were analyzed in randomly selected Basta-resistant and kanamycin-resistant T1 transgenic plants, respectively. The numbers in parentheses indicate the number of randomly selected T1 plants. Pearson correlation coefficients (R) and the coefficient of determination (R^2) for mutation rates between the target gene and the *MAR1* locus were determined for the following: The following figures present the mutation rates of the *SOS1* (**a, b**), *TOR* (**c, d**), *DME* N-terminal (**e, f**), and *DME* C-terminal (**g, h**) genes, respectively. The mutation rates of Basta-resistant (**a, c, e, g**) and kanamycin-resistant (**b, d, f, h**) T1 plants are presented. Two biological replicate experiments were conducted for *TOR* and the N- and C-termini of *DME*. **i,** Statistical analysis of mutation rates in target genes. The mutation rates of target genes were calculated following Basta (on the left) and kanamycin (on the right) screening. The standard deviation of the Student's t-test was calculated ($^*P < 0.05$, $^{***}P < 0.001$, $^{****}P < 0.0001$).

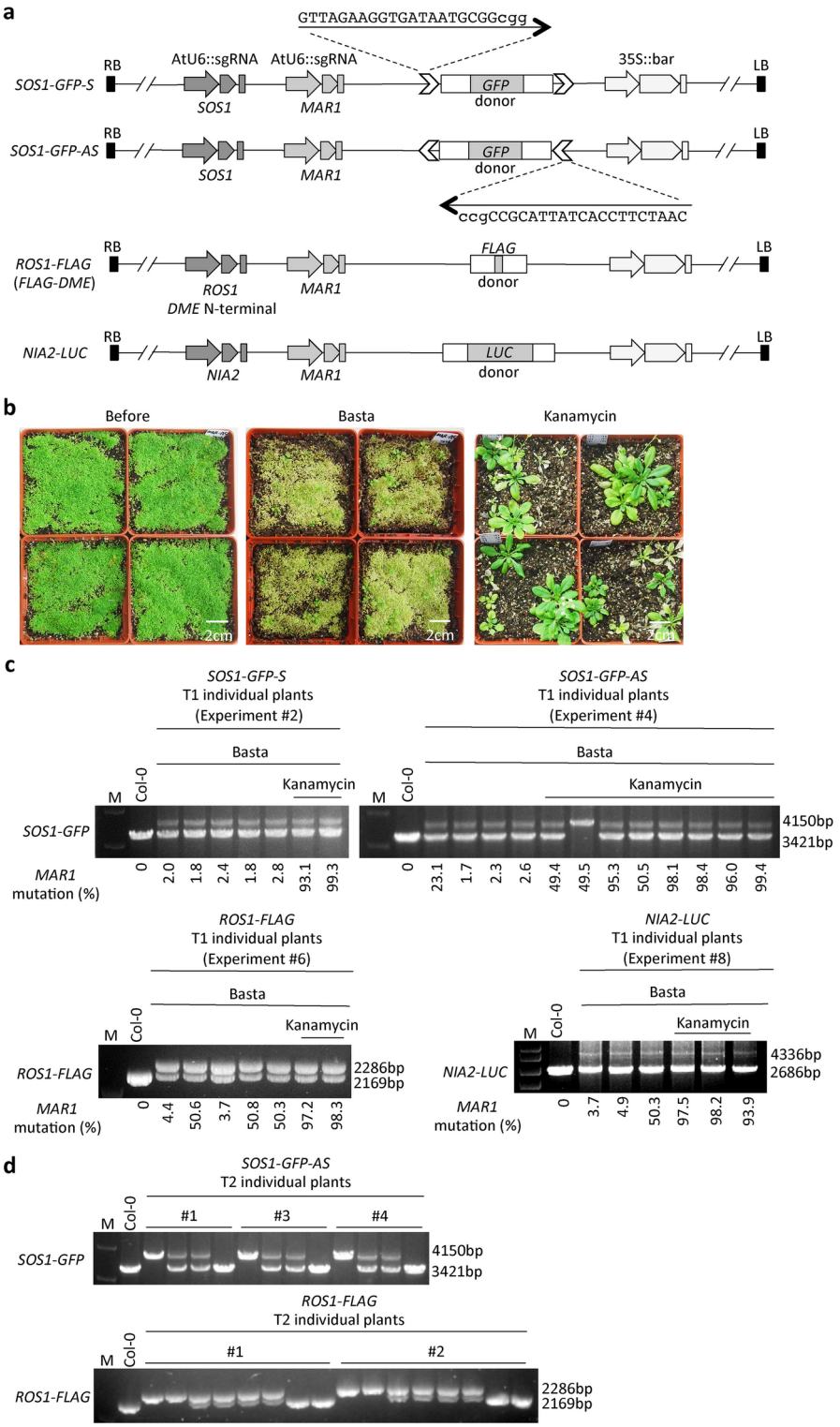

Fig. 3. Precise knock-in gene targeting via double-step screening. (**a**) A schematic representation of the KI constructs is presented. The constructs of *GFP* KI to *SOS1*, *Flag* KI to *ROS1* and N-terminus of *DME*, and *Luc* KI to *NIA2* are shown. The sgRNA recognition sites for donor excision are located at both ends of the homology arms of the *SOS1-GFP* KI constructs as sense (*SOS1-GFP-S*) or antisense (*SOS1-GFP-AS*) orientation, respectively. (**b**) Plant growth following Basta and kanamycin screening. Transgenic seeds were sown directly into the soil and screened with Basta spray followed by kanamycin. (**c**) Genotyping PCR for KI GT events. The full-length primer set was employed to identify precise KI GT events (Supplementary Fig. S1). The mutation rate at the *MAR1* locus, as determined by TIDE, is also presented. (**d**) Genotyping analysis of T2 generation. Genotypes were identified through the use of full-length primer sets.

Construct	Experiment	Basta screening	GT positive (biallelic)	GT efficiency (%)	kanamycin screening	GT positive (biallelic)	GT efficiency (%)
SOS1-GFP-S	#1	96	1	1.04	22	1	4.55
	#2	171	7	4.09	42	2	4.76
SOS1-GFP-AS	#3	112	4	3.57	28	3	10.71
	#4	237	12 (1)	5.06	52	8 (1)	15.38
ROS1-FLAG	#5	185	7	3.78	50	3	6.00
	#6	188	7	3.72	46	2	4.35
NIA2-LUC	#7	168	1	0.59	49	1	2.04
	#8	176	6	3.41	24	3	12.50
Flag-DME	#9	142	2	1.41	30	0	0

Table 1. Precise GT efficiencies GT efficiency was calculated based on the number of individual T1 transformants examined

relative GT efficiency increased by a factor of 1.1 to 3.6 (Table 1). Conversely, as with the *SOS1-GFP* system, some precise GT plants could not be selected by kanamycin screening due to the low mutation frequency at the *MAR1* locus (Fig. 3c, Table 1). Unfortunately, no precise *Flag-DME* GT plants were obtained through surrogate screening. This is due to the fact that all two precise GT plants failed to survive after kanamycin selection (Table 1). All precise GT events obtained in this study were stably inherited by progeny with Mendelian inheritance (Fig. 3d), as in previous reports[7–11].

Direct screening using surrogate endogenous marker genes

The subsequent objective was to validate whether direct screening with kanamycin would yield precise GTs in Arabidopsis. However, the results indicated that direct screening with kanamycin was not an effective method for identifying GTs in either soil or plates (Supplementary Fig. S2). A number of trials were conducted under varying screening conditions, during which numerous false positives and, potentially, false negatives were identified. A total of 36 kanamycin-resistant plants were obtained in the plate screening, and one precise GT strain was yielded from the 36 kanamycin-resistant plants with an efficiency of 2.7% (Supplementary Fig. S2). The results of direct screening with kanamycin were less clear, and even surviving plants appeared unhealthy. Furthermore, the efficacy of precise GT by direct screening with kanamycin did not exceed that of Basta-mediated transgene-based GT screening.

Furthermore, eight additional genes were examined, including *ABI5* and *WINK8*, which have been demonstrated to enhance stress tolerance in mutants with loss-of-function mutations. These phenotypes are anticipated to be valuable for surrogate screening (Supplementary Table S1). However, direct screening of the T1 generation was unsuccessful in all instances (Supplementary Fig. S3). Moreover, it seems difficult to implement the double-step screening strategy for the aforementioned eight genes, largely due to the complexity of the stress treatments. Consequently, a double-step screening approach utilizing the kanamycin-resistance phenotype in *MAR1* gene defective mutants would be a more effective method for obtaining precise GTs in Arabidopsis.

Discussion

The results demonstrate that the double-step screening strategy enables a more rigorous selection of GT candidates with high DSB frequencies, resulting in enhanced relative GT efficiency in Arabidopsis. One advantage of the double-step screening system is that the screening population can be reduced by approximately 25%, thereby improving relative GT efficiency. One disadvantage of the surrogate screening approach is that it does not guarantee the identification of all precise GT events. In conclusion, the double-step screening system appears to be a useful strategy for increasing relative GT efficiency. However, further improvements are needed to reduce the unexpected loss of precise GT plants due to low mutation frequency at the *MAR1* locus. A further disadvantage is that *mar1* mutants are unable to achieve optimal growth under conditions of iron deficiency[29]. This is due to the fact that *MAR1* plays a crucial role in the export of iron within the mitochondria and chloroplasts. In the event that physiological research is to be conducted on the generated GT plants, it would be advisable to remove the *mar1* mutations by means of backcrossing. Although unwanted mutations and randomly integrated transgenes can be removed by two rounds of backcrossing, this process is quite time-consuming[10].

It was unexpected that not only biallelic mutants but also heterozygous mutations at the *MAR1* locus with precise *SOS1-GFP* GT events were selected with kanamycin. These findings align with the results of the mutation experiments conducted in the present study, which demonstrated that both biallelic mutants and several heterozygous mutants exhibited resistance to kanamycin. This phenomenon may be attributed to the dosage of *MAR1*, as previously reported[23].

The use of surrogate systems with simultaneous editing for the purpose of efficient mutant screening has been reported by a base editor in wheat[24] and by Cas9 in Arabidopsis and tomato[22,23]. The reports suggested a positive correlation between Cas9-mediated mutations in two distinct loci that are targeted simultaneously in plants. However, to the best of our knowledge, this is the first report to demonstrate a positive correlation through the application of statistical methods. Simultaneous mutagenesis of the *MAR1* gene has been employed for efficient screening in the T0 generation in tomato, but only in the T2 generation in Arabidopsis[23]. These findings may suggest that the simultaneous editing system is not well-suited for direct screening of the first generation T1 in

Arabidopsis. This is consistent with the findings of the present study. It was found that direct screening with kanamycin and other stress treatments were unsuccessful in both soil and plate in the present study. The reason for the failure of direct screening remains uncertain. One possible explanation is that the high density of plants may have contributed to this outcome. Consequently, the double-step screening strategy, which entails initially screening with Basta to reduce plant density and then screening with kanamycin, can be employed to identify the precise GT plants with high efficiency.

It is anticipated that surrogate systems will enhance the efficiency of relative screening in GT as well as SSNs-mediated mutagenesis[13,22–24]. However, they have rarely been employed in GT in plants. During the preparation of this manuscript, a surrogate screening system for KI via NHEJ in rice was reported. In this system, a Cas9-mediated mutagenesis approach was utilized to restore resistance to hygromycin by rescuing a non-functional *HptII* gene[13]. The primary limitation of this approach is the necessity for a stable mutant version of the *HptII* gene in a transgenic plant. Another surrogate system was employed to target the *GLABRA 1* (*GL1*) gene, which is associated with the trichome-free phenotype, in Arabidopsis as a visible marker. Nevertheless, the *GL1* gene mutation phenotype was not employed for the screening of GT events in the manuscript[12].

In the aforementioned methodology, both the selection and the anticipated outcome of KI are contingent upon NHEJ[13]. In contrast, the surrogate system described in this manuscript employs NHEJ-based editing selection to detect HDR-mediated GT. It is well established that the HDR mechanism is largely cell cycle dependent, whereas NHEJ is not[4]. This leads to the question of whether the selection of GT events via HDR with NHEJ-based editing is optimal. The double-step screening method appears to have resulted in a higher relative efficiency of GT events at the three loci. This is likely due to the fact that mutagenesis at the *MAR1* locus resulted in the enrichment of plants with higher DSB activity. However, further optimization may be possible in the future.

It is noteworthy that when the sgRNA recognition sites for donor excision were arranged in the antisense direction (*SOS1-GFP-AS*), both biological replicates exhibited a slight but not statistically significant ($P=0.133$) increase in GT efficiency compared to the sense direction (*SOS1-GFP-S*) (Experiment #1 vs. #3, and #2 vs. #4). In theory, the efficiency of DSBs induced by Cas9 is not affected by the orientation of the target sequence. The underlying molecular mechanism will be subjected to further investigation.

In conclusion, the double-step screening method, which involves simultaneous mutagenesis at the endogenous marker gene *MAR1*, represents a valuable approach for enhancing the relative GT frequency in Arabidopsis up to 4.3-fold. However, it is not always effective in the present study. Given that the *MAR1* gene is highly conserved among the plant kingdom[23], the principle of the surrogate strategy established in this study would be widely applicable. One illustrative example is the biolistic approach, which involves the use of plasmids or in vitro-prepared Cas9 ribonucleoprotein complexes (RNPs) to modify the genomic sequence in crop plants. The simultaneous mutagenesis of the endogenous *MAR1* locus would be a valuable approach for efficiently screening for precise GT events in other plants. Consequently, the double-step surrogate method would be particularly advantageous for the efficient establishment and screening of GT events in plants.

Methods

Gene accession numbers

MAR1, At5g26820; *SOS1*, At2g01980; *TOR*, At1g50030; *DME*, At5g04560; *ROS1*, At2g36490; *NIA2*, At1g37130; *DD45*, At2g21740.

Plant materials and growth condition

The *Arabidopsis thaliana* (Arabidopsis) accession Col-0 and the Arabidopsis Biological Resource Center (ABRC) provided the stock number CS69955 parental line, previously named DD45-#58, which was utilized in all experiments[7]. The parental line plants of the sequential transformation GT method stably express a human codon-optimized *Streptococcus pyogenes* Cas9 in egg cells and early embryos[7]. All plants were cultivated at 22 °C on 1/2 Murashige and Skoog (MS) medium or in soil under a 16-h light/8-h dark photoperiod.

Plasmid construction

The co-editing constructs and the GT constructs for the sequential transformation strategy were constructed in accordance with the publications[7,30]. In brief, the AtU6-26 promoter-driven sgRNA cassette and donor sequence were constructed in pCambia3301. The ligation products were transformed into *E. coli*, and monoclonal bacteria were detected by colony PCR. Thereafter, recombinant plasmids were extracted and sequenced. A list of all primers utilized in this study can be found in Supplementary Table S2. The plasmid containing the sgRNA targeting the *MAR1* gene is available from Addgene (see Data availability).

Arabidopsis plant transformation

The generated constructs were transferred to *Agrobacterium tumefaciens* (Agrobacterium) GV3101 competent cells by heat shock method and spread on LB solid medium containing kanamycin and rifampicin. The cultures were incubated in the dark at 28 °C for two days to obtain positive transformants. The transformed Agrobacterium is initially cultured in 5 mL of liquid LB, subsequently grown in a 150 mL culture of LB, and finally collected by centrifugation at 4000 rpm for 20 min. The collected Agrobacterium was resuspended in an infection solution containing 5% (w/v) sucrose, 0.22% (w/v) MS, and 0.05% (v/v) Silwet-77. On the night preceding or on the day of transformation, it is advisable to remove all fruit pods and white flowers from the plants. The plant buds were then immersed in the infection solution for 45 s, removed, and gently shaken. They were then wrapped in plastic wrap to maintain humidity. The plants were placed in a dark environment at a temperature of 22 °C for a period of 20 h. The wrapping was then removed, and the plants were transferred to a normal growth environment.

T1 seeds generated by the flower dipping method were sown directly into the soil and subjected to three applications of Basta at a concentration of 0.2% (v/v), followed by five to eight applications of kanamycin at a concentration of 300 mg/L, administered every three days.

DNA analysis

Genomic DNA was extracted from leaf tissue using the cetyltrimethylammonium bromide (CTAB) method for individual plant analysis. The leaf tissues were pulverized to a fine powder in liquid nitrogen using the ShakeMaster AUTO (Bio Medical Science Inc., Tokyo, Japan). The extracted DNA was utilized for polymerase chain reaction (PCR) analysis of GT events. Specific primers were designed for genotyping purposes (see Supplementary Table S2). The PCR system utilized 2×Taq Plus Master Mix II (Vazyme, Nanjing, China), in accordance with the provided instructions. The PCR products were separated by electrophoresis on 1.5% (w/v) agarose gel and visualized using the Image Lab software (Bio-Rad Laboratories, Hercules, USA).

The TIDE website (https://tide.nki.nl) was employed to ascertain the mutation frequency of the target sites[26]. The PCR amplicons of the target sites were subjected to Sanger sequencing. A statistical analysis was performed to determine the mutation frequencies in each plant with respect to the Col-0 samples. The TIDE system provides a mutation rate range of 0 to 100%. Thus, the mutation rates are expressed on a scale from 0 to 1, with 0 indicating no mutation and 1 indicating complete mutation in the context of this study. Consequently, a mutation rate of 0 is considered to represent the wild type, 0.5 would be indicative of heterozygosity, and 1 would be indicative of biallelic.

Data availability

The plasmid bearing the AtU6-26 pro: MAR1-sgRNA sequence was deposited with Addgene and is accessible from their website (Addgene ID 210,757; pCambia3301-AtU6-MAR1). All data supporting the results of this study are presented in the manuscript, including the supplementary information. The datasets generated or analyzed during the current study are available from the corresponding author upon reasonable request.

Received: 5 June 2024; Accepted: 18 November 2024
Published online: 28 December 2024

References

1. Fauser, F. et al. In planta gene targeting. *Proc. Natl. Acad. Sci. USA* **109**, 7535–7540. https://doi.org/10.1073/pnas.1202191109 (2012).
2. Miki, D. et al. Efficient generation of diRNAs requires components in the posttranscriptional gene silencing pathway. *Sci Rep* **7**, 301. https://doi.org/10.1038/s41598-017-00374-7 (2017).
3. Gao, C. Genome engineering for crop improvement and future agriculture. *Cell* **184**, 1621–1635. https://doi.org/10.1016/j.cell.2021.01.005 (2021).
4. Miki, D. et al. Gene targeting facilitated by engineered sequence-specific nucleases: potential applications for crop improvement. *Plant Cell Physiol.* **62**, 752–765. https://doi.org/10.1093/pcp/pcab034 (2021).
5. Puchta, H. Applying CRISPR/Cas for genome engineering in plants: the best is yet to come. *Curr. Opin. Plant Biol.* **36**, 1–8. https://doi.org/10.1016/j.pbi.2016.11.011 (2017).
6. Capdeville, N., Schindele, P. & Puchta, H. Getting better all the time: recent progress in the development of CRISPR/Cas-based tools for plant genome engineering. *Curr. Opin. Biotechnol.* **79**, 102854. https://doi.org/10.1016/j.copbio.2022.102854 (2023).
7. Miki, D., Zhang, W., Zeng, W., Feng, Z. & Zhu, J.-K. CRISPR/Cas9-mediated gene targeting in Arabidopsis using sequential transformation. *Nature Commun.* **9**, 1967. https://doi.org/10.1038/s41467-018-04416-0 (2018).
8. Peng, F., Zhang, W., Zeng, W., Zhu, J.-K. & Miki, D. Gene targeting in Arabidopsis via an all-in-one strategy that uses a translational enhancer to aid Cas9 expression. *Plant Biotechnol. J.* **18**, 892–894. https://doi.org/10.1111/pbi.13265 (2020).
9. Zhang, W. et al. Precise and heritable gene targeting in rice using a sequential transformation strategy. *Cell Rep. Methods* **3**, 100389. https://doi.org/10.1016/j.crmeth.2022.100389 (2023).
10. Zhang, Z. et al. Insights into the molecular mechanisms of CRISPR/Cas9-mediated gene targeting at multiple loci in Arabidopsis. *Plant Physiol.* **190**, 2203–2216. https://doi.org/10.1093/plphys/kiac431 (2022).
11. Cheng, Y. et al. Simple promotion of Cas9 and Cas12a expression improves gene targeting via an all-in-one strategy. *Front. Plant Sci.* **15**, 1360925. https://doi.org/10.3389/fpls.2024.1360925 (2024).
12. Schreiber, T. et al. Efficient scar-free knock-ins of several kilobases in plants by engineered CRISPR-Cas endonucleases. *Mol. Plant* https://doi.org/10.1016/j.molp.2024.03.013 (2024).
13. Tian, Y. et al. High-throughput genome editing in rice with a virus-based surrogate system. *J. Integr. Plant Biol.* **65**, 646–655. https://doi.org/10.1111/jipb.13381 (2023).
14. Schindele, P. et al. Enhancing gene editing and gene targeting efficiencies in Arabidopsis thaliana by using an intron-containing version of ttLbCas12a. *Plant Biotechnol. J.* **21**, 457–459. https://doi.org/10.1111/pbi.13964 (2023).
15. Schindele, P. & Puchta, H. Engineering CRISPR/LbCas12a for highly efficient, temperature-tolerant plant gene editing. *Plant Biotechnol. J.* **18**, 1118–1120. https://doi.org/10.1111/pbi.13275 (2020).
16. Li, J., Kong, D., Ke, Y., Zeng, W. & Miki, D. Application of multiple sgRNAs boosts efficiency of CRISPR/Cas9-mediated gene targeting in Arabidopsis. *BMC Biol.* **22**, 6. https://doi.org/10.1186/s12915-024-01810-7 (2024).
17. Li, J. et al. Cas12a-mediated gene targeting by sequential transformation strategy in Arabidopsis thaliana. *BMC Plant Biol.* **24**, 665. https://doi.org/10.1186/s12870-024-05375-z (2024).
18. Merker, L., Schindele, P., Huang, T.-K., Wolter, F. & Puchta, H. Enhancing in planta gene targeting efficiencies in Arabidopsis using temperature-tolerant CRISPR/LbCas12a. *Plant Biotechnol. J.* **18**, 2382–2384. https://doi.org/10.1111/pbi.13426 (2020).
19. Wolter, F. & Puchta, H. In planta gene targeting can be enhanced by the use of CRISPR/Cas12a. *Plant J.* **100**, 1083–1094. https://doi.org/10.1111/tpj.14488 (2019).
20. Zhang, H. et al. A surrogate reporter system for multiplexable evaluation of CRISPR/Cas9 in targeted mutagenesis. *Sci. Rep.* **8**, 1042. https://doi.org/10.1038/s41598-018-19317-x (2018).
21. Jung, S. B., Lee, C. Y., Lee, K. H., Heo, K. & Choi, S. H. A cleavage-based surrogate reporter for the evaluation of CRISPR-Cas9 cleavage efficiency. *Nucleic Acids Res.* **49**, e85. https://doi.org/10.1093/nar/gkab467 (2021).
22. Kong, X. et al. GLABRA2-based selection efficiently enriches Cas9-generated nonchimeric mutants in the T1 generation. *Plant Physiol.* **187**, 758–768. https://doi.org/10.1093/plphys/kiab356 (2021).
23. Rinne, J., Witte, C.-P. & Herde, M. Loss of MAR1 function is a marker for co-selection of Crispr-induced mutations in plants. *Front Genome Ed.* **3**, 723384. https://doi.org/10.3389/fgeed.2021.723384 (2021).

24. Zhang, R. et al. Generation of herbicide tolerance traits and a new selectable marker in wheat using base editing. *Nature plants* **5**, 480–485. https://doi.org/10.1038/s41477-019-0405-0 (2019).
25. Aufsatz, W. et al. A novel strategy for obtaining kanamycin resistance in Arabidopsis thaliana by silencing an endogenous gene encoding a putative chloroplast transporter. *Biotechnol. J.* **4**, 224–229. https://doi.org/10.1002/biot.200800156 (2009).
26. Brinkman, E. K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42**, e168–e168. https://doi.org/10.1093/nar/gku936 (2014).
27. Dickinson, L. et al. Regulation of gene editing using T-DNA concatenation. *Nature Plants* **9**, 1398–1408. https://doi.org/10.1038/s41477-023-01495-w (2023).
28. Merker, L., Feller, L., Dorn, A. & Puchta, H. Deficiency of both classical and alternative end-joining pathways leads to a synergistic defect in double-strand break repair but not to an increase in homology-dependent gene targeting in Arabidopsis. *Plant J.* **118**, 242–254. https://doi.org/10.1111/tpj.16604 (2024).
29. Kim, L. J. et al. Ferroportin 3 is a dual-targeted mitochondrial/chloroplast iron exporter necessary for iron homeostasis in Arabidopsis. *Plant J.* **107**, 215–236. https://doi.org/10.1111/tpj.15286 (2021).
30. Miki, D. *et al.* in *Arabidopsis Protocols Fourth Edition, Methods in Molecular Biology* (eds Jose J. Sanchez-Serrano & Julio Salinas) pp. 121–146 (Springer, 2021)

Acknowledgements

This work was supported by the Shanghai Science and Technology Innovation Plan (20ZR1467000 and 23WZ2500800), the Foreign Expert Project (G202201355L), and the Chinese Academy of Sciences to D.M. We would like to thank all lab members of Epigenetics and Genome engineering group and the Shanghai Center for Plant Stress Biology, CAS Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences for assistance.

Author contributions

Y.Q.C., L.Z., and D.M. designed the research; Y.Q.C., L.Z., and D.M. performed the experiments with assistance from Y.P.K., and X.F.D.; Y.Q.C., L.Z., and D.M. wrote the paper. All authors contributed to the article and approved the submitted version.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-80352-y.

Correspondence and requests for materials should be addressed to D.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.