

ORIGINAL RESEARCH

American Society
of Plant Biologists
Cultivating a better future through plant biology research

WILEY

Multiplex genome editing in *Arabidopsis thaliana* using Mb3Cas12a

William T. Jordan | Seth Currie | Robert J. Schmitz

Department of Genetics, University of Georgia, Athens, Georgia, USA

CorrespondenceRobert J. Schmitz, University of Georgia,
Department of Genetics, 120 East Green
Street, Athens, GA 30602, USA.
Email: schmitz@uga.edu**Funding information**UGA Office of Research; Pew Charitable
Trusts; National Science Foundation, Grant/
Award Number: MCB-1856143; National
Institute of General Medical Sciences of the
National Institutes of Health, Grant/Award
Number: T32GM007103**Abstract**

The use of CRISPR-Cas proteins for the creation of multiplex genome engineering represents an important avenue for crop improvement, and further improvements for creation of knock-in plant lines via CRISPR-based technologies may enable the high-throughput creation of designer alleles. To circumvent limitations of the commonly used CRISPR-Cas9 system for multiplex genome engineering, we explored the use of *Moraxella bovoculi* 3 Cas12a (Mb3Cas12a) for multiplex genome editing in *Arabidopsis thaliana*. We identified optimized cis-regulatory sequences for driving expression of single-transcript multiplex crRNA arrays in *A. thaliana*, resulting in stable germline transmission of Mb3Cas12a-edited alleles at multiple target sites. By utilizing this system, we demonstrate single-transcript multiplexed genome engineering using of up to 13 crRNA targets. We further show high target specificity of Mb3Cas12a-based genome editing via whole-genome sequencing. Taken together, our method provides a simplified platform for efficient multiplex genome engineering in plant-based systems.

KEYWORDS

Arabidopsis, Cas12a, Cpf1, CRISPR, gene editing, Mb3Cas12a

1 | INTRODUCTION

The use of bacterial Cas nucleases for genome editing has enabled facile manipulation of numerous plant genomes, including *Arabidopsis thaliana*, maize, tomato, soybean, and rice (Brooks et al., 2014; Jacobs et al., 2015; Jiang et al., 2013; Li et al., 2013; Xing et al., 2014). Cas enzymes, such as Cas9, function as DNA endonucleases which recognize their cognate DNA by complementary base pairing with a tracrRNA and target-specific crRNA. For genome-editing purposes in eukaryotic organisms, the tracrRNA and crRNA are often fused into a single chimeric guide RNA (gRNA) (Jinek et al., 2012). Cas9 from *Streptococcus pyogenes* (SpyCas9) is by far the most widely adopted enzyme currently used for plant genome editing, due to its high activity at room temperature in numerous plant species, acceptable

rates of off-targeting, widespread availability in numerous plant-compatible vectors, and the availability of several engineered variants (Brooks et al., 2014; Li et al., 2013). Through the use of sequence-specific chimeric single-guide RNA scaffold (sgRNA), sequences immediately adjacent to the protospacer adjacent motif (PAM) of 5' NGG 3' can be targeted for DNA binding and blunt-end cleavage (Jinek et al., 2012).

As sgRNAs must match their DNA targets with sequence specificity, Pol-III promoters are typically required to express sgRNAs, due to the lack of 5' and 3' processing of Pol-III transcripts (Jacobs et al., 2015; Xing et al., 2014). However, Pol-III promoters suffer from a lack of tissue and environmental specificity, lack of nuclear export of transcripts, and contain cryptic termination poly(T) termination sequences, which are present within the sgRNA scaffold and could

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. *Plant Direct* published by American Society of Plant Biologists and the Society for Experimental Biology and John Wiley & Sons Ltd.

limit expression of sgRNA transcripts (Ui-Tei et al., 2017). Additionally, while the 3-bp PAM of Cas9 occurs on average once every 16 bp, finding suitable gRNA sequences in AT-rich regions such as promoters and some genic regions can be challenging.

Due to the inherent lack of RNase activity in Cas9 nucleases, approaches for targeting of multiple genomic regions using Cas9 (multiplex genome editing) require the use of promoter stacking for individual gRNA cassettes or the recruitment and/or introduction of sequence-specific RNA processing machinery. While promoter stacking is commonly utilized in plants for multiplex genome editing, construction of stacked constructs requires numerous cloning steps and frequently utilizes a small number of repeated promoter sequences, resulting in repetitive clones that are prone to bacterial recombination.

Alternatives to promoter stacking using Pol-III promoters include the insertion of RNA sequences that result in RNA cleavage at desired locations, either through the recruitment of endogenous RNases or self-cleavage mechanisms. Successful recruitment of endogenous RNases for gRNA processing is routinely accomplished by the introduction of tRNAs flanking gRNA cassettes, which serve to facilitate processing by RNase P and RNase Z. Alternatively, ribozyme sequences flanking gRNA cassettes can also be used, albeit with lower reported efficiency when compared with tRNA-based gRNA processing (Čermák et al., 2017).

The use of exogenous RNA processing machinery, such as CRISPR-associated RNA endoribonuclease 4 (Csy4), has also been successfully adapted to process gRNA cassettes from a single-transcript molecule in plants, with the addition of a flanking 20-bp Csy4 recognition sequence (Čermák et al., 2017). While the use of exogenous RNases requires expression of an additional protein, the use of Csy4 to process polycistronic gRNAs has been demonstrated to outperform the use of ribozymes and native tRNA processing machinery for obtaining genome-edited plants (Čermák et al., 2017).

While Cas9 remains the most utilized CRISPR-Cas nuclease for genome-editing purposes, an alternative CRISPR-Cas-based system for genome editing, CRISPR-Cas12a (formerly called Cpf1), has been utilized for genome-editing purposes in animals, plants, and fungi. Cas12a enzymes employ unique mechanisms for RNA processing, DNA cleavage, and distinct PAM sequences (Zetsche et al., 2015). Cas12a enzymes requiring only a short 20-bp crRNA scaffold for target recognition and can self-process multiple crRNAs from a single molecule, through the utilization of an RNase domain for specifically processing crRNA arrays (Zetsche et al., 2016). Cas12a enzymes recognize a T-rich PAM (typically 5' TTTV 3'), enabling easier targeting of AT-rich genomic regions when compared with Cas9, and additionally produce staggered cuts with a 5-bp overhang, often resulting in the formation of larger indels when compared with Cas9-based genome editing (Zetsche et al., 2015). Studies examining off-target genome editing have shown reduced off-target activity with Cas12a enzymes in mammalian systems when compared with Cas9, a result further supported in studies involving rice (Kleinstiver et al., 2016; Tang et al., 2018; Zhang et al., 2018).

Several previous studies have shown the efficacy of LbCas12a-based gene editing in *A. thaliana* by introducing single or dual crRNAs per gene target using a double-ribozyme system (Bernabé-Orts et al., 2019; Malzahn et al., 2019). However, the successful application of alternative Cas12a orthologs, as well as the use of Cas12a for massively multiplexed genome editing, remains unexplored in plants. Here, we report the use of Mb3Cas12a for genome editing in plants and demonstrate the one-step introduction of up to 13 crRNAs in a single ORF for massively multiplexed genome editing. When coupled with a modified heat-shock protocol previously described in plants (LeBlanc et al., 2018), we report the isolation of homozygous T2 populations with up to six different edits at unique target sites from a single T1 parent plant.

2 | RESULTS

2.1 | Mb3Cas12a is a functional RNase and DNase in *A. thaliana*

To test the functionality of a human codon-optimized Mb3Cas12a in plant systems, we constructed binary vectors based on the pK11.1R vector backbone (Figure 1a–c). The pK11.1R backbone was selected due to the high efficiency of the *RIBOSOMAL PROTEIN 5A* (*RPS5A*) promoter for Cas9-based mutagenesis, which was used to drive expression of Mb3Cas12a for all experiments. For crRNA array expression, the efficiency of Pol-II and Pol-III promoters was compared by expressing identical crRNA arrays under either the U6-26 Pol-III promoter or a Pol-II promoter. To identify suitable Pol-II promoters for crRNA expression, RNA-seq data from the shoot apical meristem (SAM) at various developmental timepoints were analyzed to identify candidate genes with the highest expression across developmental time (Klepikova et al., 2015). The expression of *TRANSLATIONALLY CONTROLLED TUMOR PROTEIN1* (*AtTCTP1*; AT3G16640) possessed high, consistent expression across developmental time in the SAM, with an average transcripts per million (TPM) count of 3157. This was much higher in comparison with expression of *RPS5A* (TPM = 1309) and *POLYUBIQUITIN 10* (*UBI10*, TPM = 325). Additionally, multiple *AtTCTP1* upstream *cis*-regulatory fragments and tissue specificity of transgenic expression genes driven using the promoter of *AtTCTP1* have been previously experimentally validated (Han et al., 2015). To assess the ability of Mb3Cas12a for genome editing, a four-crRNA array consisting of two guides targeting the floral regulator genes *APATELLA1* (*AP1*; AT1G69120) and two targeting *CAULIFLOWER* (*CAL*; AT1G26310) using both *pU6-26* and *pAtTCTP1* was transformed via floral dip and T1 transformants screened via hygromycin selection and grown at 22°C with a 16-h photoperiod. A total of 64 and 63 plants containing Pol-III-driven or Pol-II-driven *AP1*-*CAL* arrays were analyzed for the presence of mutant phenotypes and genotypes, respectively. Previous studies have showed that mutations in *ap1* result in homeotic conversion of sepals to bracts and petals are converted to stamens, whereas mutation in *cal* result has no obvious phenotypes. However, a double

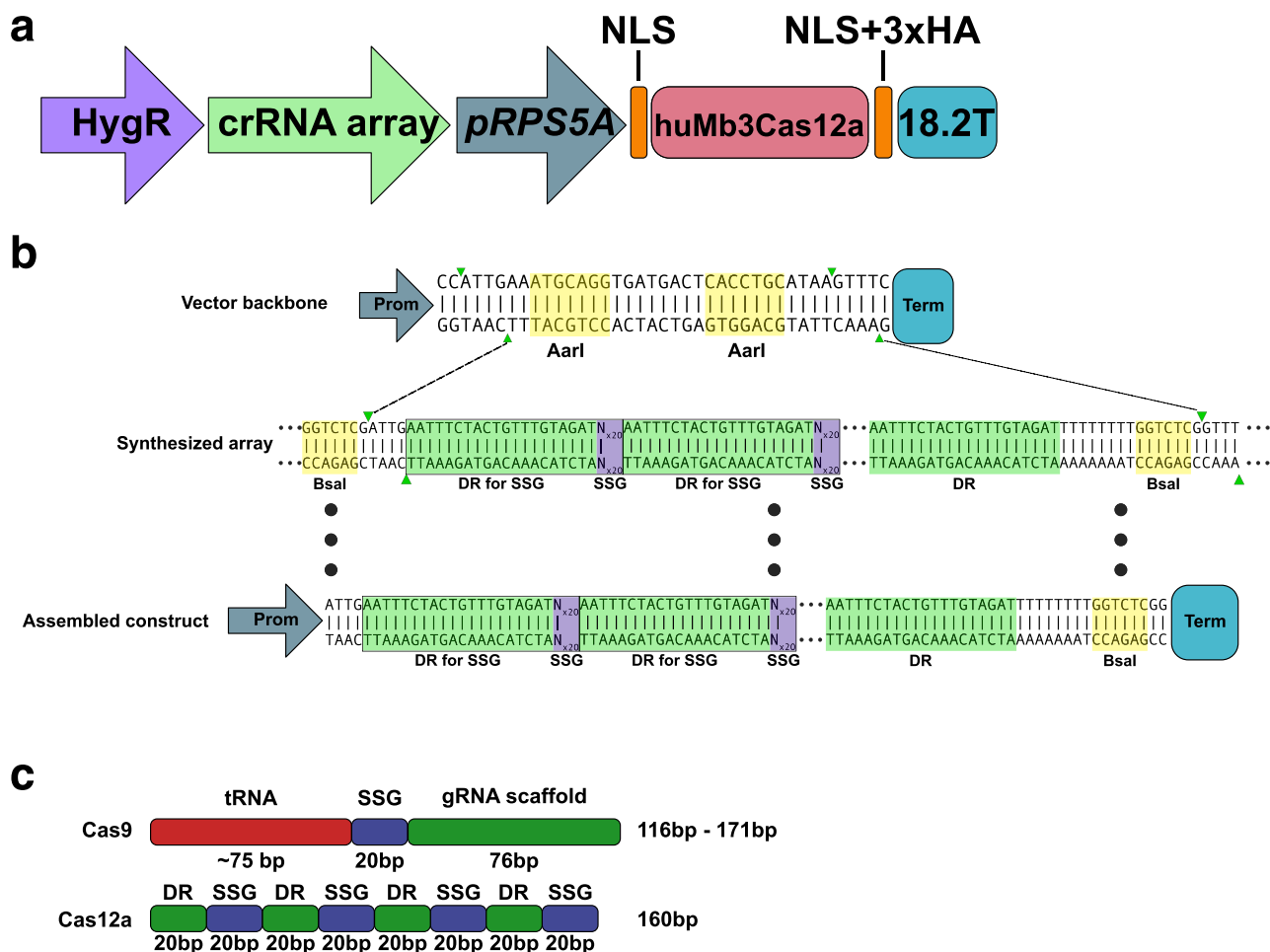


FIGURE 1 Construction and cloning protocol of Mb3Cas12a vectors. (a) Vector design consisting of a hygromycin resistance cassette, crRNA expression cassette, and huMb3Cas12a driven by the *AtRPS5A* promoter. NLS = nuclear localization signal; 18.2 T = Hsp18.2 terminator from *Arabidopsis thaliana*. (b) Cloning strategy for insertion of custom crRNA arrays. Custom crRNA arrays are cloned in a single step using Golden-Gate Assembly via *AarI* restriction sites. DR = Mb3Cas12a direct repeat sequence; SSG = sequence-specific guide. (c) Size comparison (bp) of SSG cassettes for Cas9 and Cas12a-based multiplexing systems in plants

mutation of *ap1;cal* results in a very strong phenotype of an inflorescence that resembles a cauliflower. Unexpectedly, no mutant phenotypes or genotypes were detected for both Pol-II and Pol-III transformed lines.

To investigate if temperature-sensitive activity of Mb3Cas12a was resulting in undetectable editing activity, T2 seeds from a single T1 transformant was split into three groups and subject to three heat treatments (Figure 2a), and the emergence of expected phenotypes was observed upon transition to reproductive growth. Plants subjected to a 30°C heat stress during the entirety of vegetative growth, in addition to individuals grown at room temperature, failed to exhibit any observable mutant genotypes and phenotypes, while approximately 10% (1/10) of T2 plants transformed with a Pol-II-driven array subject to repeated 37°C heat shocks exhibited observable *ap1* mutant phenotypes and genotypes (Figure 2b,c). The observed molecular genotypes and phenotypes were stable for at least one generation post heat treatment, indicating that 37°C heat treatment is sufficient for detectable and inherited edits for Pol-II-driven arrays (Figure 2d).

2.2 | Assessment of editing of inflorescences using a single crRNA array

To investigate the efficiency and specificity of multiplex genome editing, the floral regulator genes *APETALA 2* (*AP2*), *PISTILLATA* (*PI*), and *AGAMOUS* (*AG*) were selected for editing using a single Pol-II array. Successful editing of these genes in all floral meristem cells would induce distinct phenotypes depending on the number and types of mutations found in these three targeted genes. For example, a triple mutant results in all leaf-like whorls, whereas in a *pi ag* mutant, whorls are converted to sepals, in *ap2 pi* mutants, whorls are converted to carpels, and in *ap2 ag* mutants results in leaves in the first and fourth whorls and modified petals/stamens in the second and third whorls. Each of these mutants also has distinct mutant phenotypes in isolation. The *ag* mutants have homeotic conversion of stamens to petals and carpels to sepals and flowers. The *ap2* mutant has homeotic conversion of sepals and petals, and *pi* mutants have abnormal petal and stamen identity. Importantly, the proportion and

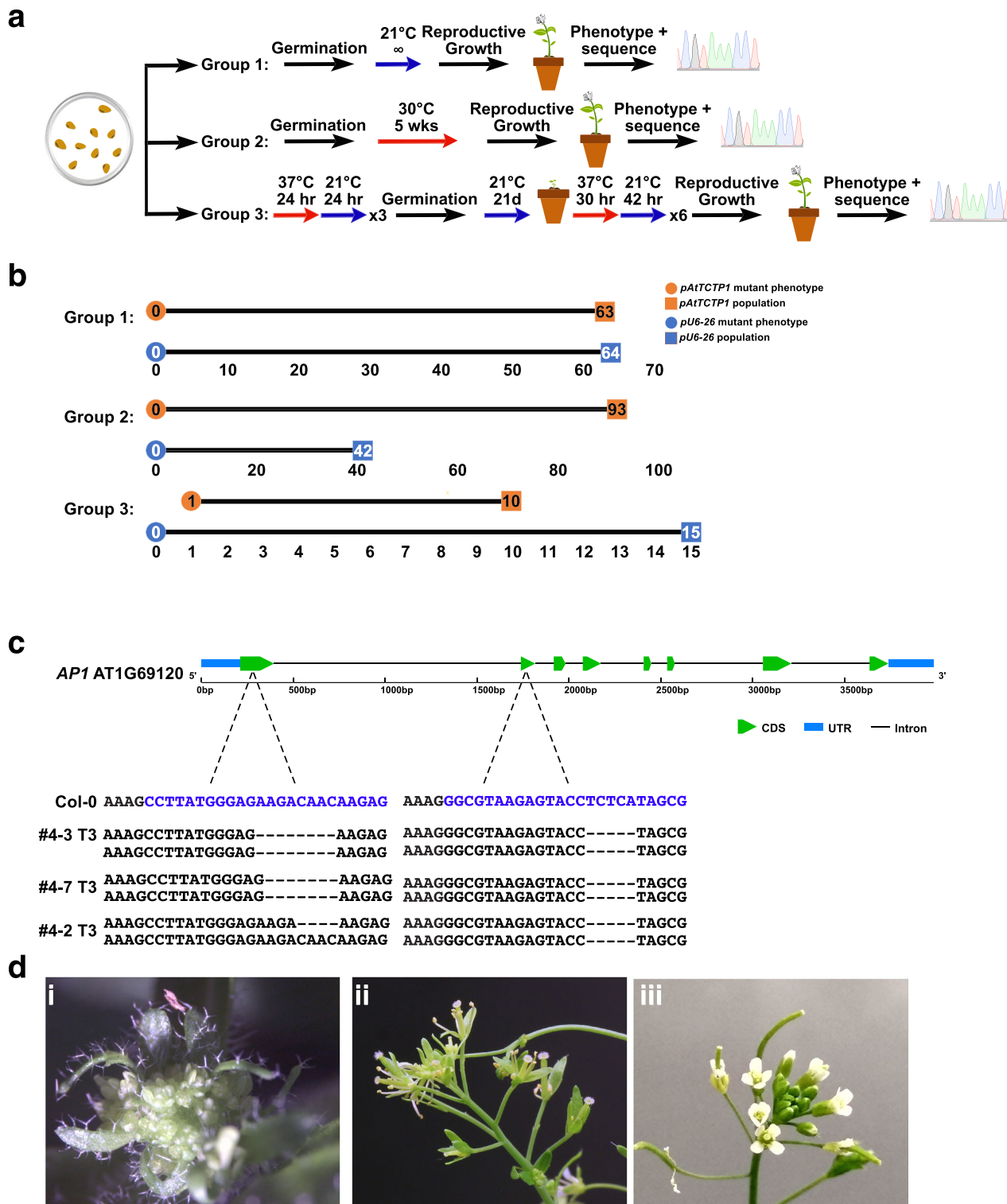


FIGURE 2 Assessment of Mb3Cas12a activity at different growth temperatures. (a) Identical T2 seed pools were subjected to three temperature conditions: constant growth at 21°C (1), growth at 30°C until flowering was observed (2), or repeated heat shocks at 37°C (3). (b) Assessment of mutant phenotypes observed in T2 pools using either *pU6-26* or *pAtTCTP1* for crRNA array expression targeting *AP1* and *CAL* using various temperature conditions. Pools with mutant phenotypes were saved and propagated an additional generation to obtain stable T3 lines. (c) Sanger sequencing of T3 lines reveal deletions at crRNA target sites. PAM sites are indicated by black text and the target site is indicated by blue text in the Col-0 sequence. (d) Edited lines resemble *ap1* with respect to inflorescence meristem architecture (i) and floral architecture (ii) compared with Col-0 wild type (iii)

identification of heterozygous genotypes in T1 plants can be ascertained by phenotyping T2 progeny, as homozygous or biallelic mutations in *PI* and *AG* result in sterility.

Four unique guides were used to target the *AP2* and *PI* loci, whereas two overlapping guides containing a single mismatch at Positions 13 and 18 were used to target the *AG* locus (Figure 3a). Twelve independent T1 plants were subjected to 37°C heat shocks (Figure 2a), and the emergence of floral phenotypes were assessed upon transition to reproductive growth. Upon flowering, three plants exhibited sectoring with altered floral morphology, mimicking previously described *ap2* and *pi* mutants, as well as *ap2 pi* double mutants (Figure 3b). Interestingly, no *ag*-like sectors were observed on any T1 plants, indicating reduced editing activity of single

guide mismatches at the *ag* locus when compared with the *AP2* and *PI* loci.

The inheritance of induced mutations following 37°C heat treatments from these three lines were subsequently assessed by collecting and growing T2 seeds at a constant temperature of 22°C, to eliminate potential additional Mb3Cas12a endonuclease activity. Upon flowering, 16% to 38% of T2 plants exhibited *ap2*-like flower phenotypes (Figure 3b). Sanger sequencing performed at the *AP2* locus revealed homozygous deletions at a single crRNA target site ranging from −4 to −18 bp, supporting previous observations that Cas12a-mediated genome editing primarily results in the formation of deletions between 5 and 15 bp in size (Malzahn et al., 2019). Despite the greatly reduced fertility of *ap2* mutants, homozygous mutations

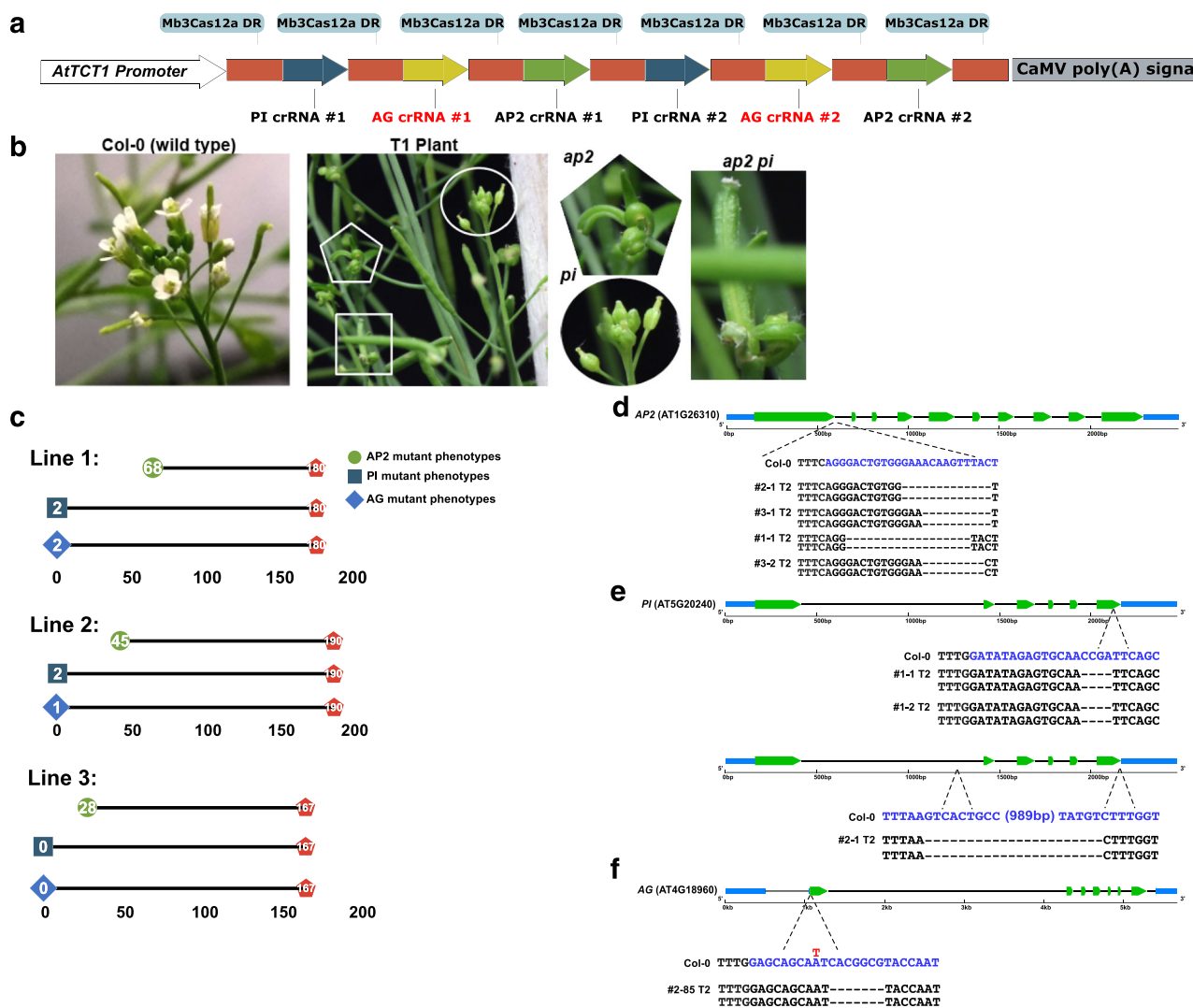


FIGURE 3 Assessment of multiplex gene editing using Mb3Cas12a. (a) Schematic of crRNA array targeting three floral regulator genes (*APETALA2*, *PISTILLATA*, *AGAMOUS*) expressed from the *AtTCTP1* promoter. Guides targeting *AG* contain a SNP at positions 19 and 13, respectively. (b) Sectoring observed within T1 individuals mimicking single and higher order floral regulator mutations in comparison with Col-0 wild type. (c) Assessment of mutant phenotypes observed in T2 pools from T1 lines exhibiting floral phenotypes. A lack of *PI* knockout phenotypes in the T2 likely represents low levels of heterozygosity from T1-edited plants. (d–f) Sanger sequencing of T2 lines exhibiting floral defects at *AP2*, *PI*, and *AG* loci. Only the crRNA target site with observed edits is shown. PAM sites are indicated by black text and the target site is indicated by blue text in the Col-0 sequence



were readily observed in T2 populations, indicating a high degree of inheritance of induced mutations, despite negative selective pressure.

To ascertain the level of heterozygosity in T1 parent plants, T2 populations were assessed for editing at the *PI* and *AG* loci. As *pi* mutants are sterile due to the lack of anther formation, and *ag* mutants fail to produce anthers and pistils, homozygous *pi* and/or *ag* mutants observed in T2 individuals would result from residual heterozygosity in their respective T1 parents. Out of 180 and 190 observed T2 plants for Lines 1 and 2, only two individuals for each line were observed to have homozygous or biallelic mutations at the *PI* locus for each line, indicating a low degree of heterozygosity observed with 37°C heat treatments. Interestingly, Sanger sequencing of the *PI* locus revealed deletions ranging from −4 to −1003 bp in *pi* T2 individuals, indicating Mb3Cas12a can induce a large range of insertion/deletion (indel) sizes in *A. thaliana*. A similar number of *ag* phenotypes were observed in T2 populations, indicating a similar rate of transmission of heterozygous *AG* and *PI* alleles in T1 plants. As *ag* mutant sectors were not observed in T1 individuals, the rate of Mb3Cas12a-induced editing at crRNA sites containing single mismatches is substantially reduced.

2.3 | Efficacy of noncanonical PAM editing using Mb3Cas12a

Previous characterization of Mb3Cas12a activity in mammalian cell culture systems revealed the ability to recognize alternate or non-canonical PAM sequences for genome editing. To determine the efficiency of noncanonical PAM sequences for editing in *A. thaliana*, four crRNAs targeting the *SHOOT APICAL MERISTEM ARREST 1* (*SHA1*) locus with either TTN or CTN PAM motifs were transformed and T1 individuals subject to 37°C heat shocks (Figure 2a). Of 23 and 13 plants observed with TTN and CTN PAM arrays, respectively, no detectable editing of the *SHA1* locus was observed in T1 individuals based on screening for *sha1* mutant phenotypes. To assess the potential of editing in the SAM in hemizygous T1 individuals, 15 seeds from each T1 transformant were grown at a constant temperature of 22°C. No detectable editing was observed based on a lack of *sha1* mutant phenotypes from any T2 individuals sampled, indicating low non-canonical PAM activity of Mb3Cas12a in *A. thaliana* at the *SHA1* locus. While no activity was observed at the *SHA1* locus, testing of noncanonical PAM editing at additional target sites is necessary to confirm low noncanonical PAM activity in *A. thaliana*. Future efforts including evaluating an expanded range of target sites and inspection via sequencing will be required to determine the efficacy of non-canonical PAM editing by Mb3Cas12a.

2.4 | Multiplexed Mb3Cas12a mutagenesis and off-target analysis using Mb3Cas12a

Previous studies conducted in plants have used CRISPR-Cas mutagenesis to rapidly engineer *cis*-regulatory element variation via

Cas9-multiplexed gRNA systems (Rodriguez-Leal et al., 2017). To investigate the use of Mb3Cas12a for multiplexed mutagenesis of *cis*-regulatory regions, a crRNA array targeting 830 bp of the *FLOWERING LOCUS T* (*FT*) enhancer BlockE was constructed (Figure 4a,b). To maximize the likelihood of successful editing at multiple target sites, crRNAs were subject to target efficiency calculations using the CRISPR-DT webtool (Zhu & Liang, 2019). Thirteen crRNAs with a predicted efficiency of greater than 0.8 were selected and cloned as a single dsDNA block. Nineteen independent T1 lines were subjected to 37°C heat shocks as previously described, and editing efficiency was assessed by sequencing pools of T2 individuals from individual lines grown at a constant temperature of 22°C. Upon sequencing, three pools contained edits at the BlockE region. Of lines containing edits, approximately 60% of sequenced individuals contained edits at least one single crRNA site. In total, editing was detected at five unique crRNA targets, with deletions detected ranging in size from 6 to 581 bp (Figure 4c).

To assess potential off-target effects from the introduction of multiple crRNAs, whole-genome sequencing (WGS) was performed on seven T2 BlockE individuals from two independently transformed lines at an average sequencing depth of ~41× coverage ranging from ~19× to ~62× coverage. Putative off-target deletions were identified using DeepVariant 1.0. A total of 329 high-confidence deletions were detected across all sequenced individuals, with at least 96% of detected deletions shared between all individuals, indicating shared deletions inherited from T0 parents (Figure 5a–c). To further investigate potential sources of unique deletions between samples, the size of uniquely occurring deletions was plotted (Figure 5c). Over 80% of the detected deletions were outside the observed range of previously observed Cas12a-induced indels (deletions between 3 and 25 bases), indicating a majority of detected deletions are likely Cas12a-independent polymorphisms. To assess the possibility of Cas12a off-target editing at deletions between 3 and 25 bp in size, these variants were further filtered for proximity to a Cas12a PAM site. All variants containing a 3- to 25-bp deletion and within 20 bp of a Cas12a PAM site were observed to be located in the BlockE region, indicating no off-target Cas12a deletions observed in the seven single individuals sequenced and supporting previous studies indicating the high specificity of Cas12a-mediated genome editing in eukaryotic systems (Strohkendl et al., 2018).

3 | DISCUSSION

In this study, we created and optimized a simplified CRISPR-Cas genome-editing system for multiplex targeting, using the Mb3Cas12a nuclease. As previous studies have indicated temperature sensitivity for Cas12a-mediated DNase activity, we first investigated which temperature profile would enable efficient editing in *A. thaliana* (Kleinstiver et al., 2016). Our results show no detectable editing at temperatures of 21°C and 30°C. Interestingly, previous studies have shown activity of LbCas12a at 30°C in *A. thaliana* and even 22°C in rice, suggesting that temperature sensitivity for plant genome editing

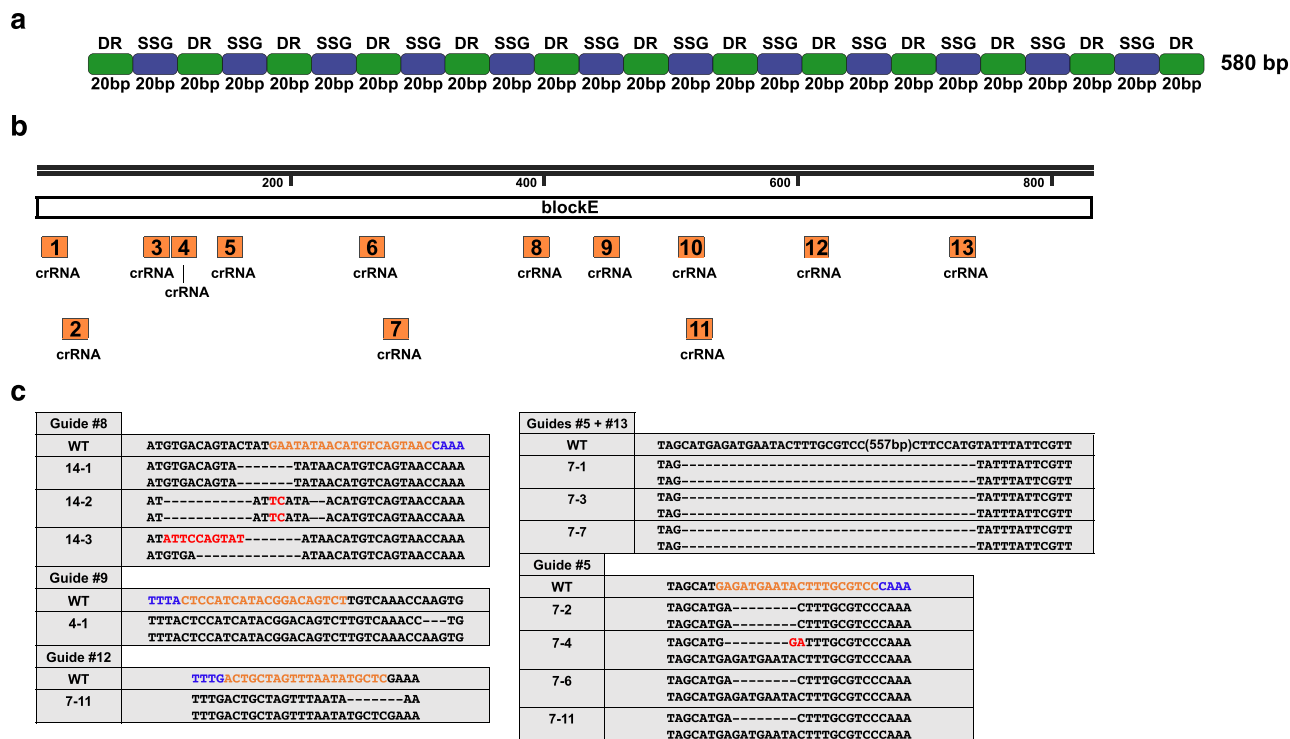


FIGURE 4 CRE mutagenesis using Mb3Cas12a. (a) Schematic of a 13crRNA array targeting the *FT* regulatory element BlockE. (b) Location of high-quality crRNAs span the entire BlockE region of *FT* and allow for recovery of numerous edited alleles. (c) Edited alleles obtained from T2 transgene-free individuals. Edits were observed at six target sites, resulting in indel patterns ranging from −3 to −587 bp. PAM sites are indicated by blue text and the target site is indicated by orange text in the Col-0 sequence

varies widely depending on the Cas12a ortholog and plant systems used (Malzahn et al., 2019). When applying a modified heat-shock protocol as previously described in LeBlanc et al. (2018), we observed potent levels of editing in recovered somatic tissue, indicating the minimum temperature for DNase activity of Mb3Cas12a is between 31°C and 37°C.

Using a U6-driven crRNA array, we first tested the ability of Mb3Cas12a to properly process multiple crRNAs from a single RNA transcript, by creating mutations at two distinct sites at the *AP1* locus. Although we were able to observe somatic mutations in rosette leaf samples, poor germline transmission of edited alleles was observed using this system. As previous studies in mammalian systems indicated increased mutagenesis efficiency using Cas12a when crRNAs were driven by a Pol-II promoter, we performed an exhaustive search for promoters with high expression levels, specifically in the SAM of *A. thaliana* (Klepikova et al., 2015; Zhong et al., 2017). Upon the utilization of the 0.3-kb *AtTCTP1* promoter, we observed genome editing in both somatic and germline tissue of T2 plants. As *AtTCTP1* is one of the most highly expressed genes in dividing tissues, and particularly the SAM, it is likely the increased germline mutagenesis observed when using a *pAtTCTP1*-driven crRNA expression system is due to increased abundance of crRNA molecules available for processing in the SAM.

We next investigated the efficiency of targeting multiple loci simultaneously using a Pol-II-driven array. Upon flowering, we

observed both single and double mutants targeting the *AP2* and *PI* loci. T2 progeny from these edited plants exhibited high levels of homozygous editing of the *AP2* locus, indicating efficient germline transmission of edited alleles, despite reduced fitness of *ap2* flowers. Interestingly, we observed relatively few *pi* and *ap2 pi* T2 plants. As *pi* mutants are sterile, this observation supports biallelic or homozygous editing in T1 plants, which were unable to be transmitted through the germline due to sterility. To test the effects of single mismatches on editing rates, we simultaneously introduced two mismatched guides with a single mismatch targeting the *AG* locus. When compared with nonmismatched guides targeting *AP2*, editing rates at *AG* were 50-fold lower, indicating there could be a severe penalty for mutagenesis using singly mismatched crRNAs. Alternatively, the sequence context of the specific target site was not as susceptible to genome editing.

To test the ability of massively multiplex editing of genomic regions, using a one-step cloning approach, we introduced 13 crRNAs in a single array for targeting the BlockE region, adjacent to the *FT* locus into *A. thaliana*. BlockE has previously been shown to function as a downstream enhancer of *FT*, and artificial deposition of DNA methylation at BlockE has resulted in a delayed flowering phenotype (Zicola et al., 2019). T2 plants recovered were observed to have deletions at numerous target sites, ranging in size from 3 to 587 bp. Interestingly, no delayed flowering phenotypes in BlockE-edited lines when grown in long-day conditions were observed. As the

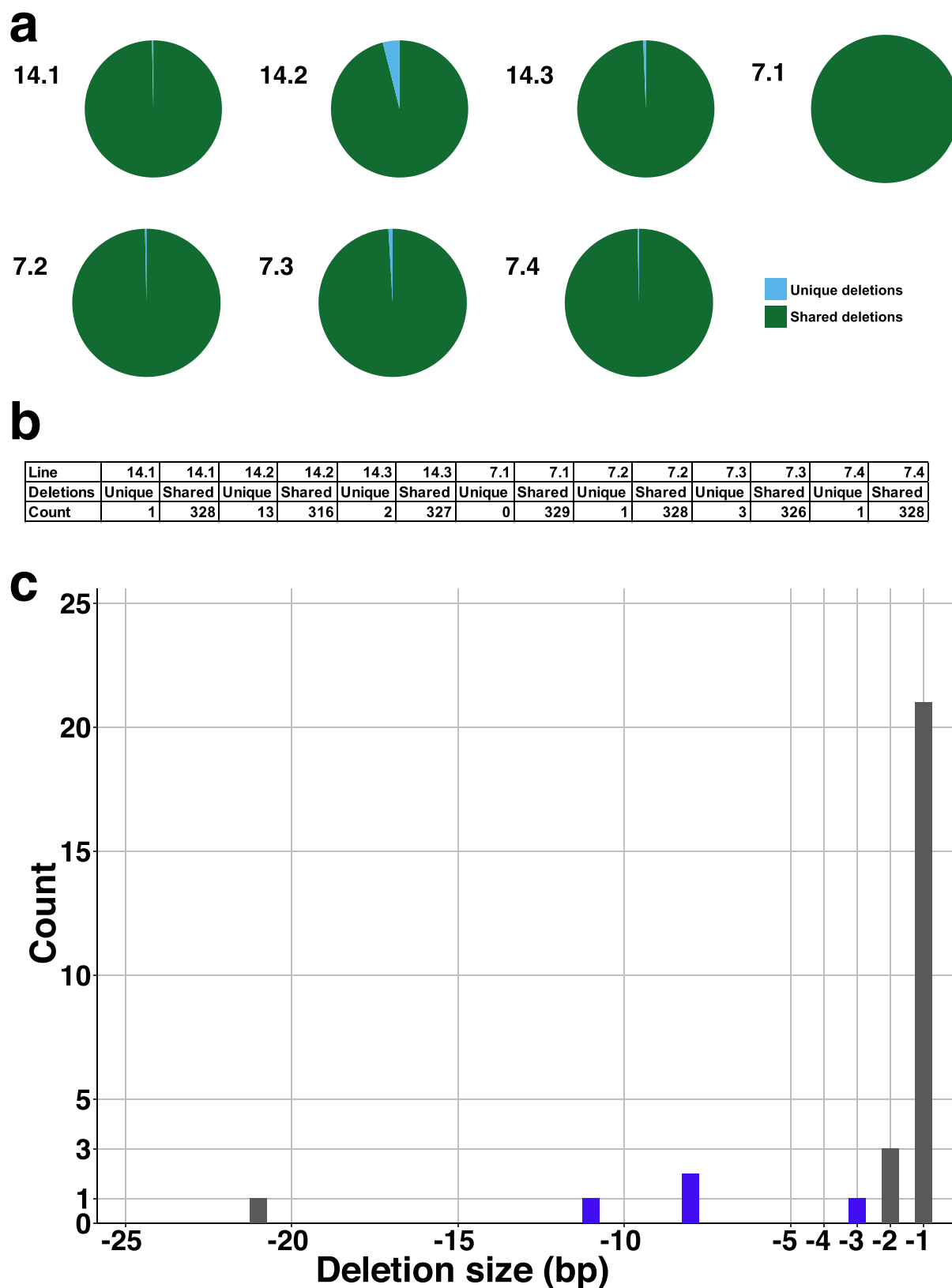


FIGURE 5 Detected variants in BlockE-targeted T2 individuals. (a) Number of unique and shared deletions in BlockE L7 and L14 T2 individuals. (b) Count of unique and shared deletions in BlockE L7 and L14 T2 individuals. (c) Size distribution of nonshared deletions in all BlockE individuals



introduction of large numbers of crRNAs could potentially increase off-targeting rates, we profiled the presence of off-target editing in seven T2 plants using WGS. As no off-target edits were detected in this population of edited plants, the introduction of multiple crRNAs with predicted target specificity is unlikely to drastically affect off-targeting rates genome wide.

The simplicity of construction of multiplex targeting using our system, combined with the efficiency of inducing mutations when using a high-temperature heat-shock protocol, enables a simplified, efficient massively multiplexed mutagenesis in *A. thaliana*. Importantly, this occurs without the introduction or recruitment of endogenous RNases or exogenous RNA processing proteins. Future efforts to bypass the need for heat shock by introducing candidate temperature-tolerant Cas12a variants identified in a previous study would be a worthwhile pursuit (Merker et al., 2020). Lastly, by reducing the size and repeat content of transgenic material needed for targeting of Cas proteins to multiple genomic sites, this Mb3Cas12a expression system enables more facile multiplex editing in plant-based systems.

4 | METHODS

4.1 | Plant transformation, selection, and heat treatment

Vectors were transformed into *Agrobacterium tumefaciens* strain C58C1 by electroporation, and transformation of Col-0 *A. thaliana* inflorescence tissue was performed using the floral dip method (Clough & Bent, 1998). Seeds were collected upon plant senescence, and transgenic plants were identified via selection on 1/2 strength LS plates supplemented with hygromycin B (25 µg/ml) as described previously (Harrison et al., 2006). Approximately 10–14 days following germination, resistant seeds were selected and planted into SunGro #3B Mix. Ten days following transplanting, plants were subjected to eight rounds of heat stress as described previously with minor modifications. Plants were subsequently recovered at room temperature during reproductive growth and grown at 21°C with a 16-h photoperiod.

4.2 | Cloning

The ORF containing NLS-Mb3Cas12a-3xHA was subcloned from 35-pcDNA3-huMb3Cpf1 (a gift from Feng Zhang) into the pKI1.1R backbone by GenScript, creating Mb3Cas12a-pKI1.1R. To construct pAtTCTP1-Mb3Cas12a-pKI1.1R, 1 µg of Mb3Cas12a-pKI1.1R was doubly digested using 50 units of Apa I (New England Biolabs) and 2 units of Eco RI-HF (New England Biolabs). The native 0.3-kb AtTCTP1 promoter and CaMV poly(A) signal cassette was synthesized as a single FragmentGENE DNA fragment (Genewiz), doubly digested with Apa I and Eco RI and ligated to the Mb3Cas12a-pKI1.1R using 3000 units of T7 DNA Ligase (New England Biolabs). crRNA arrays were cloned by amplifying synthesized

FragmentGENE or PriorityGene (Genewiz) fragments with primers crRNA F (5' GTAGTCGTAGTCGGTCTC 3') and crRNA R (5' GGACTCCGTGGATACAAA 3') using Q5 DNA Polymerase (New England Biolabs). The resulting amplicons were cleaned using a Monarch PCR & DNA Cleanup Kit (New England Biolabs). Approximately 300 ng of cleaned PCR product were digested with 12 units of BsaI-HFv2 (New England Biolabs), gel purified using a Monarch PCR & DNA Cleanup Kit and inserted into Aar I-digested Mb3Cas12a-pKI1.1R with T7 DNA Ligase (New England Biolabs), and transformed into DH10B cells using electroporation. All clones were sequence verified using Sanger sequencing.

4.3 | DNA extraction and analysis

DNA for Sanger sequencing analysis was extracted from rosette leaves using previously described method (Edwards et al., 1991). Candidate regions were amplified with Q5 DNA Polymerase and sequenced using Sanger sequencing (Macrogen, USA). Mutated alleles were deconvoluted from Sanger sequencing traces using CRISP-ID (Dehairs et al., 2016).

4.4 | DNA extraction and WGS analysis

DNA for Illumina NGS analysis was extracted using a DNeasy Plant Mini Kit (Qiagen) following the manufacturer's instructions. Genomic DNA library preparation was carried out using the protocol previously described (Ji et al., 2018) and paired-end 150-bp reads were sequenced using an Illumina NovaSeq 6000. Raw fastq reads were preprocessed using fastp v0.20.1, aligned to the TAIR10 genome using Bowtie 2.3.5.1 and duplicates removed using Sambamba v.6.6. Variants were called using DeepVariant v1.0, and gVCF files merged and joint variants were called using GLnexus v1.2.7 (Poplin et al., 2018). Obtained merged gVCF files were filtered for deletion-specific variants using Jvarkit v20200206, and sites containing sufficient data for all sequenced lines and a minimum sequencing depth of 10 were filtered using VCFtools v0.1.16.

ACKNOWLEDGMENTS

We would like to thank Bernd Zetsche, Rhiannon Macrae, and Feng Zhang for sharing the Mb3Cas12a plasmid for use in plants. W.T.J. was supported by National Institute of General Medical Sciences of the National Institutes of Health award number T32GM007103. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This study was funded by the support from the National Science Foundation (MCB-1856143), the Pew Charitable Trusts and the UGA Office of Research to R.J.S.

CONFLICT OF INTEREST

The authors declare no conflict of interest associated with the work described in this manuscript.



AUTHOR CONTRIBUTIONS

W.T.J. and R.J.S. designed research. W.T.J. and S.C. performed research. W.T.J. analyzed data and wrote the paper.

DATA AVAILABILITY STATEMENT

All raw and processed whole-genome sequencing data have been deposited to NCBI GEO and can be accessed using accession number GSE162875.

ORCID

Seth Currie  <https://orcid.org/0000-0002-0047-1545>

Robert J. Schmitz  <https://orcid.org/0000-0001-7538-6663>

REFERENCES

- Bernabé-Orts, J. M., Casas-Rodrigo, I., Minguet, E. G., Landolfi, V., García-Carpintero, V., Gianoglio, S., Vázquez-Vilar, M., Granell, A., & Orzaez, D. (2019). Assessment of Cas12a-mediated gene editing efficiency in plants. *Plant Biotechnology Journal*, 17, 1971–1984. <https://doi.org/10.1111/pbi.13113>
- Brooks, C., Nekrasov, V., Lippman, Z. B., & Van Eck, J. (2014). Efficient gene editing in tomato in the first generation using the clustered regularly interspaced short palindromic repeats/CRISPR-associated9 system. *Plant Physiology*, 166, 1292–1297. <https://doi.org/10.1104/pp.114.247577>
- Čermák, T., Curtin, S. J., Gil-Humanes, J., Čegan, R., Kono, T. J. Y., Konečná, E., Belanto, J. J., Starker, C. G., Mathre, J. W., Greenstein, R. L., & Voytas, D. F. (2017). A multipurpose toolkit to enable advanced genome engineering in plants. *The Plant Cell*, 29, 1196–1217. <https://doi.org/10.1105/tpc.16.00922>
- Clough, S. J., & Bent, A. F. (1998). Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal*, 16, 735–743. <https://doi.org/10.1046/j.1365-313x.1998.00343.x>
- Dehairs, J., Talebi, A., Cherifi, Y., & Swinnen, J. V. (2016). CRISP-ID: Decoding CRISPR mediated indels by Sanger sequencing. *Scientific Reports*, 6, 28973. <https://doi.org/10.1038/srep28973>
- Edwards, K., Johnstone, C., & Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research*, 19, 1349–1349. <https://doi.org/10.1093/nar/19.6.1349>
- Han, Y. J., Kim, Y. M., Hwang, O. J., & Kim, J. I. (2015). Characterization of a small constitutive promoter from *Arabidopsis* translationally controlled tumor protein (AtTCTP) gene for plant transformation. *Plant Cell Reports*, 34, 265–275. <https://doi.org/10.1007/s00299-014-1705-5>
- Harrison, S. J., Mott, E. K., Parsley, K., Aspinall, S., Gray, J. C., & Cottage, A. (2006). A rapid and robust method of identifying transformed *Arabidopsis thaliana* seedlings following floral dip transformation. *Plant Methods*, 2, 1–7. <https://doi.org/10.1186/1746-4811-2-19>
- Jacobs, T. B., LaFayette, P. R., Schmitz, R. J., & Parrott, W. A. (2015). Targeted genome modifications in soybean with CRISPR/Cas9. *BMC Biotechnology*, 15, 1–10. <https://doi.org/10.1186/s12896-015-0131-2>
- Ji, L., Jordan, W. T., Shi, X., Hu, L., He, C., & Schmitz, R. J. (2018). TET-mediated epimutagenesis of the *Arabidopsis thaliana* methylome. *Nature Communications*, 9, 1–9. <https://doi.org/10.1038/s41467-018-03289-7>
- Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., & Weeks, D. P. (2013). Demonstration of CRISPR/Cas9/sgrRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. *Nucleic Acids Research*, 41, e188. <https://doi.org/10.1093/nar/gkt780>
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337, 816–821. <https://doi.org/10.1126/science.1225829>
- Kleinstiver, B. P., Tsai, S. Q., Prew, M. S., Nguyen, N. T., Welch, M. M., Lopez, J. M., McCaw, Z. R., Aryee, M. J., & Joung, J. K. (2016). Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nature Biotechnology*, 34, 869–874. <https://doi.org/10.1038/nbt.3620>
- Klepikova, A. V., Logacheva, M. D., Dmitriev, S. E., & Penin, A. A. (2015). RNA-seq analysis of an apical meristem time series reveals a critical point in *Arabidopsis thaliana* flower initiation. *BMC Genomics*, 16, 1–16. <https://doi.org/10.1186/s12864-015-1688-9>
- LeBlanc, C., Zhang, F., Mendez, J., Lozano, Y., Chatpar, K., Irish, V. F., & Jacob, Y. (2018). Increased efficiency of targeted mutagenesis by CRISPR/Cas9 in plants using heat stress. *The Plant Journal*, 93, 377–386. <https://doi.org/10.1111/tbj.13782>
- Li, J.-F., Norville, J. E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G. M., & Sheen, J. (2013). Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nature Biotechnology*, 31, 688–691. <https://doi.org/10.1038/nbt.2654>
- Malzahn, A. A., Tang, X., Lee, K., Ren, Q., Sretenovic, S., Zhang, Y., Chen, H., Kang, M., Bao, Y., Zheng, X., Deng, K., Zhang, T., Salcedo, V., Wang, K., Zhang, Y., & Qi, Y. (2019). Application of CRISPR-Cas12a temperature sensitivity for improved genome editing in rice, maize, and *Arabidopsis*. *BMC Biology*, 17, 1–14. <https://doi.org/10.1186/s12915-019-0629-5>
- Merker, L., Schindele, P., Huang, T., Wolter, F., & Puchta, H. (2020). Enhancing in planta gene targeting efficiencies in *Arabidopsis* using temperature-tolerant CRISPR/LbCas12a. *Plant Biotechnology Journal*, 18, 2382–2384. <https://doi.org/10.1111/pbi.13426>
- Poplin, R., Chang, P. C., Alexander, D., Schwartz, S., Colthurst, T., Ku, A., Newburger, D., Dijamco, J., Nguyen, N., Afshar, P. T., Gross, S. S., Dorfman, L., McLean, C. Y., & DePristo, M. A. (2018). A universal SNP and small-indel variant caller using deep neural networks. *Nature Biotechnology*, 36, 983–987. <https://doi.org/10.1038/nbt.4235>
- Rodríguez-Leal, D., Lemmon, Z. H., Man, J., Bartlett, M. E., & Lippman, Z. B. (2017). Engineering quantitative trait variation for crop improvement by genome editing. *Cell*, 171, 470–480. <https://doi.org/10.1016/j.cell.2017.08.030>
- Strohkendl, I., Saifuddin, F. A., Rybarski, J. R., Finkelstein, I. J., & Russell, R. (2018). Kinetic basis for DNA target specificity of CRISPR-Cas12a. *Molecular Cell*, 71, 816–824. <https://doi.org/10.1016/j.molcel.2018.06.043>
- Tang, X., Liu, G., Zhou, J., Ren, Q., You, Q., Tian, L., Xin, X., Zhong, Z., Liu, B., Zheng, X., Zhang, D., Malzahn, A., Gong, Z., Qi, Y., Zhang, T., & Zhang, Y. (2018). A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. *Genome Biology*, 19, 1–13. <https://doi.org/10.1186/s13059-018-1458-5>
- Ui-Tei, K., Maruyama, S., & Nakano, Y. (2017). Enhancement of single guide RNA transcription for efficient CRISPR/Cas-based genomic engineering. *Genome*, 60, 537–545. <https://doi.org/10.1139/gen-2016-0127>
- Xing, H.-L., Dong, L., Wang, Z.-P., Zhang, H.-Y., Han, C.-Y., Liu, B., Wang, X.-C., & Chen, Q.-J. (2014). A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biology*, 14, 1–12. <https://doi.org/10.1186/s12870-014-0327-y>
- Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., Volz, S. E., Joung, J., van der Oost, J., Regev, A., Koonin, E. V., & Zhang, F. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, 163, 759–771. <https://doi.org/10.1016/j.cell.2015.09.038>



- Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E. M., Winblad, N., Choudhury, S. R., Abudayyeh, O. O., Gootenberg, J. S., Wu, W. Y., Scott, D. A., Severinov, K., van der Oost, J., & Zhang, F. (2016). Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nature Biotechnology*, 35, 31–34.
- Zhang, Q., Xing, H.-L., Wang, Z.-P., Zhang, H.-Y., Yang, F., Wang, X.-C., & Chen, Q.-J. (2018). Potential high-frequency off-target mutagenesis induced by CRISPR/Cas9 in *Arabidopsis* and its prevention. *Plant Molecular Biology*, 96, 445–456. <https://doi.org/10.1007/s11103-018-0709-x>
- Zhong, G., Wang, H., Li, Y., Tran, M. H., & Farzan, M. (2017). Cpf1 proteins excise CRISPR RNAs from mRNA transcripts in mammalian cells. *Nature Chemical Biology*, 13, 839–841. <https://doi.org/10.1038/nchembio.2410>
- Zhu, H., & Liang, C. (2019). CRISPR-DT: Designing gRNAs for the CRISPR-Cpf1 system with improved target efficiency and specificity. *Bioinformatics*, 35, 2783–2789. <https://doi.org/10.1093/bioinformatics/bty1061>
- Zicola, J., Liu, L., Tanzler, P., & Turck, F. (2019). Targeted DNA methylation represses two enhancers of FLOWERING LOCUS T in *Arabidopsis thaliana*. *Nature Plants*, 5, 300–307. <https://doi.org/10.1038/s41477-019-0375-2>

How to cite this article: Jordan, W. T., Currie, S., & Schmitz, R. J. (2021). Multiplex genome editing in *Arabidopsis thaliana* using Mb3Cas12a. *Plant Direct*, 5(9), e344. <https://doi.org/10.1002/pld3.344>