

High-efficiency genome editing in plants mediated by a Cas9 gene containing multiple introns

Ramona Grützner¹, Patrick Martin², Claudia Horn¹, Samuel Mortensen³, Erin J. Cram³, Carolyn W.T. Lee-Parsons^{4,5}, Johannes Stuttmann² and Sylvestre Marillonnet^{1,*}

¹Department of Cell and Metabolic Biology, Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle (Saale), Germany

²Institute for Biology, Department of Plant Genetics, Martin Luther University Halle-Wittenberg, Weinbergweg 10, 06120 Halle (Saale), Germany

³Department of Biology, Northeastern University, Boston, MA, USA

⁴Department of Chemistry and Chemical Biology, Northeastern University, Boston, MA, USA

⁵Department of Chemical Engineering, Northeastern University, Boston, MA, USA

*Correspondence: Sylvestre Marillonnet (smarillo@ipb-halle.de)

<https://doi.org/10.1016/j.xplc.2020.100135>

ABSTRACT

The recent discovery of the mode of action of the CRISPR/Cas9 system has provided biologists with a useful tool for generating site-specific mutations in genes of interest. In plants, site-targeted mutations are usually obtained by the stable transformation of a Cas9 expression construct into the plant genome. The efficiency of introducing mutations in genes of interest can vary considerably depending on the specific features of the constructs, including the source and nature of the promoters and terminators used for the expression of the Cas9 gene and the guide RNA, and the sequence of the Cas9 nuclease itself. To optimize the efficiency of the Cas9 nuclease in generating mutations in target genes in *Arabidopsis thaliana*, we investigated several features of its nucleotide and/or amino acid sequence, including the codon usage, the number of nuclear localization signals (NLSs), and the presence or absence of introns. We found that the Cas9 gene codon usage had some effect on its activity and that two NLSs worked better than one. However, the highest efficiency of the constructs was achieved by the addition of 13 introns into the Cas9 coding sequence, which dramatically improved the editing efficiency of the constructs. None of the primary transformants obtained with a Cas9 gene lacking introns displayed a knockout mutant phenotype, whereas between 70% and 100% of the primary transformants generated with the intronized Cas9 gene displayed mutant phenotypes. The intronized Cas9 gene was also found to be effective in other plants such as *Nicotiana benthamiana* and *Catharanthus roseus*.

Keywords: CRISPR, Cas9, targeted mutagenesis, gene targeting

Grützner R., Martin P., Horn C., Mortensen S., Cram E.J., Lee-Parsons C.W.T., Stuttmann J., and Marillonnet S. (2021). High-efficiency genome editing in plants mediated by a Cas9 gene containing multiple introns. Plant Comm. 2, 100135.

INTRODUCTION

The discovery of the mode of action of the CRISPR/Cas9 system in 2012 has provided biologists with a useful tool for generating site-specific mutations in genes of interest in living organisms (Jinek et al., 2012; Cong et al., 2013; Mali et al., 2013). The enzymes of several CRISPR systems, including Cas9 and Cas12/Cpf1, have since been used for generating site-targeted mutations in a wide range of organisms, including animals, plants, fungi, and bacteria. In plants, site-specific mutations are generally obtained by transforming plants with a construct that contains the Cas9 gene and guide RNA expression cassette. Cells from the transformed plants expressing all components of the CRISPR/Cas9 system undergo cleavage of chromosomal DNA at the target site(s). The cleaved DNA is then usually repaired

by non-homologous end joining, resulting in mutated target-site sequences. The Cas9 construct can be eliminated from the genome of the transformed plants in the next generation by Mendelian segregation. In plants, like in other organisms, not all transformants containing the Cas9 gene and guide RNA construct display mutations in the target gene, with efficiencies reported to vary from a few percent to close to 100% (Belhaj et al., 2013; Ahmad et al., 2020). The efficiency with which mutations are generated in target genes depends on multiple factors, including the choice of target sites in selected genes, as well as

Published by the Plant Communications Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

the nature of the coding and regulatory sequences of the Cas9 gene and guide RNA construct. In recent studies, several architectural parameters of Cas9 constructs designed for plants have been investigated, including the codon usage of the Cas9 gene, the number of nuclear localization signals (NLSs) in the Cas9 enzyme, the nature of the promoters and terminators of the Cas9 gene, the length and sequence of the conserved region of the guide RNA and the terminator sequence of the guide RNA, and the relative orientations of the various expression cassettes in the final transfer DNA (T-DNA) (Castel et al., 2019; Ordon et al., 2019). These studies suggested that several features of the construct, such as the nature of the promoters that drive Cas9 expression and the relative orientations of the guide RNAs and Cas9 gene cassettes, had some effects on the efficiency of the constructs. However, the most significant effect on efficiency was attributed to the nature of the coding sequence of the Cas9 gene. (Castel et al., 2019) compared four different Cas9 coding sequences and found that the most efficient one consisted of a plant codon-optimized sequence that contained one intron and two NLSs (the Cas9 gene initially published by Li et al., 2013). However, the plant codon-optimized sequence was not tested without an intron and was tested only with two NLSs; thus, it was not possible to conclude how much each feature individually contributed to the high activity of this particular sequence.

In this work, we investigated the role of the codon usage, the numbers of NLSs, and the presence of introns in the coding sequence of the Cas9 gene on genome editing efficiency in stable transgenic plants. Introns have been known for many years to have a positive effect on gene expression (Callis et al., 1987). For example, the introduction of multiple introns into a tobacco mosaic virus viral vector construct transiently delivered to *Nicotiana benthamiana* leaves as a T-DNA by *Agrobacterium* led to a large increase in the expression of a recombinant protein from the viral vector (Marillonnet et al., 2005). The improvement increased with the number of introns introduced, with the highest efficiency obtained with 12 introns. This beneficial effect of introns was hypothesized to be due to the increased efficiency of processing and export of pre-mRNA viral vector transcripts from the nucleus to the cytoplasm of the host cells.

Here, we found that the presence of two NLSs in Cas9 improves the efficiency of generating mutations, but the greatest effect was obtained by the addition of 13 introns to the coding sequence of the Cas9 gene. With the intronized Cas9 gene, single and double knockouts or large chromosomal deletions (30–70 kb) were obtained with high frequencies (>70% and ~10%, respectively) in primary *Arabidopsis* transformants. The intronized Cas9 gene was shown to also work well in *N. benthamiana* and *Catharanthus roseus*.

RESULTS

Comparison of Cas9 genes

Seven Cas9 expression constructs were prepared with different Cas9 coding sequences. The first Cas9 coding sequence (hCas9, Figure 1 and Supplemental Figure 1) is a human codon-optimized sequence (Mali et al., 2013). This sequence contains a C-terminal NLS. We used the domesticated version of this gene (lacking internal type IIS restriction sites; Castel

et al., 2019; Nekrasov et al., 2013) in order to be able to subclone it with the modular cloning system MoClo (Weber et al., 2011). This Cas9 sequence was subcloned without or with an additional N-terminal NLS to generate Cas9 expression constructs pAGM51511 and pAGM51613, respectively. To check the effect of codon optimization, a second Cas9 coding sequence was synthesized using the *Zea mays* (*Z. mays*) codon usage (zCas9), which has a high GC content of 55% (<https://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=4577>). High-GC-content codon usage was selected as we planned to introduce introns into this sequence, and GC-rich exon sequences can induce more efficient splicing of the inserted introns (Carle-Urioste et al., 1997). The codon-optimized Cas9 sequence contains a C-terminal NLS with the same amino acid sequence as the hCas9 gene. This Cas9 sequence was also subcloned without or with an additional N-terminal NLS to generate Cas9 expression constructs pAGM51523 and pAGM51535, respectively. A third version of the Cas9 sequence (zCas9i) was prepared by introducing 13 *Arabidopsis* introns into the *Z. mays* codon-optimized version. This Cas9 sequence was also subcloned without or with an additional N-terminal NLS to generate two Cas9 expression constructs pAGM51547 and pAGM51559, respectively. Finally, a fourth Cas9 version similar to the intron-containing Cas9, zCas9io, was prepared. It contains mutations at four sites in introns 1, 3, 12, and 13 to remove a potential cryptic splice site in the first intron and to improve the splicing efficiency of some weak splice sites, as predicted by NetGene2 intron splice site prediction software (<http://www.cbs.dtu.dk/services/NetGene2/>) (Supplemental Figure 2). This Cas9 version also contains an N-terminal FLAG tag followed by an NLS. It was subcloned to prepare the Cas9 expression construct pAGM51561. The Cas9 coding sequences in all constructs were driven by the *Arabidopsis* RPS5a promoter, which effectively expresses the Cas9 nuclease in *Arabidopsis* (Tsutsui and Higashiyama, 2017; Ordon et al., 2019).

To test different Cas9 genes, we used a single guide RNA (sgRNA) that targets the genes encoding two homologous *Arabidopsis* transcription factors, TRY and CPC. These genes act as negative regulators of trichome development and serve as convenient visual markers of Cas9 activity, with knockout plants displaying an increased number of trichomes in leaves (Wang et al., 2015). The guide RNA was cloned under the control of an *Arabidopsis* U6 promoter (Nekrasov et al., 2013) (sequence in Supplemental Figure 14).

The constructs were used to transform *Arabidopsis* using flower dip transformation, and more than 160 transformants were obtained for each construct. The constructs containing Cas9 without introns did not produce a single primary transformant with a complete knockout phenotype. A few plants displayed some leaf sectors with the expected mutant hairy phenotype (Figure 1; Supplemental Figure 3). The highest number of plants with a chimeric phenotype was obtained with the *Z. mays* codon-optimized Cas9 containing two NLSs (6/182 transformants or 3.2%). The other constructs carrying Cas9 without introns led to approximately 1% of plants with a chimeric mutant phenotype (7/664), with all other plants displaying a wild-type (WT) phenotype. By contrast, transformation of the intronized Cas9 constructs led to a large number of primary transformants with full knockout phenotypes (Figure 1; Supplemental Figure 3).

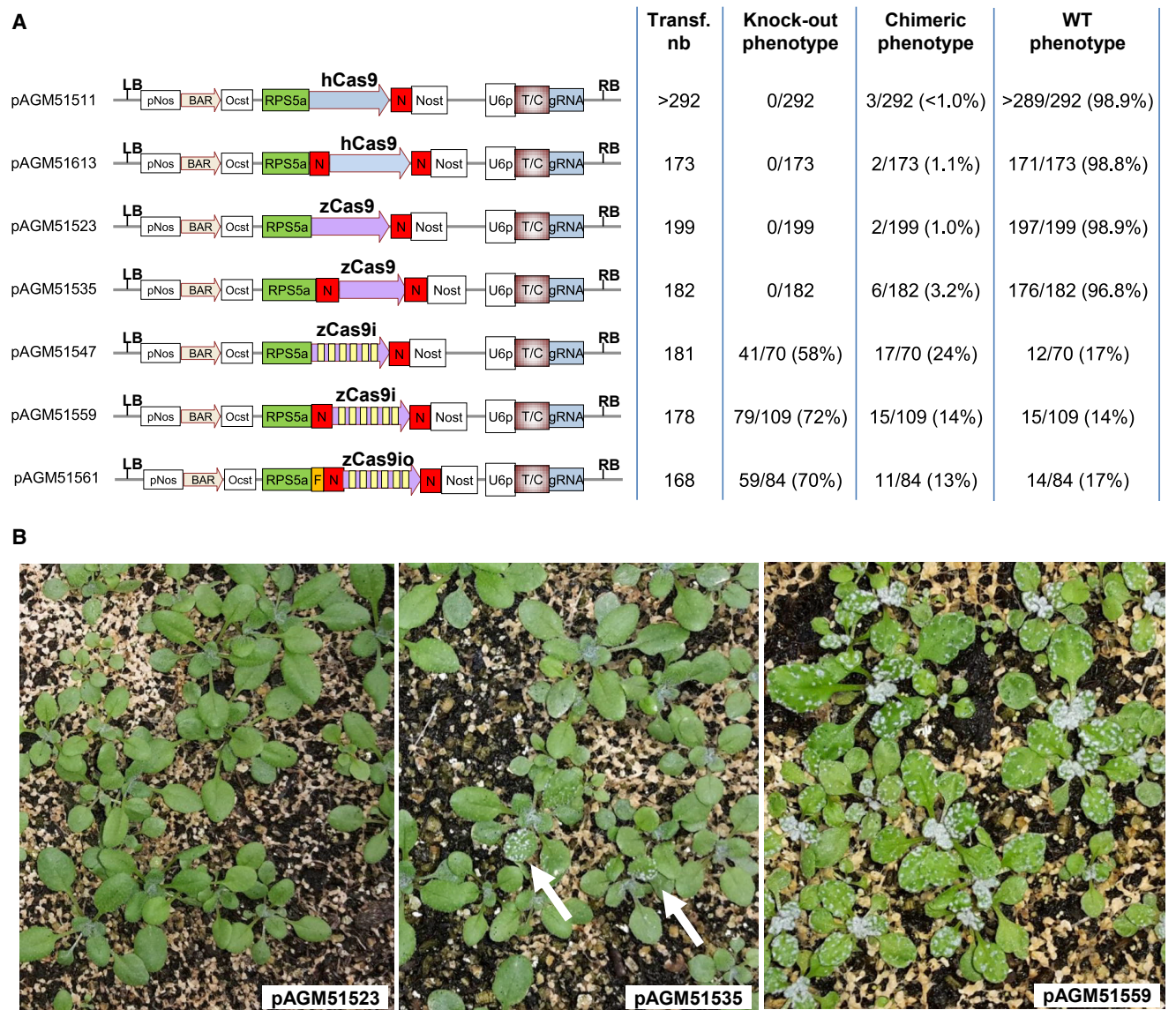


Figure 1. Comparison of different Cas9 versions by the mutagenesis of *TRY* and *CPC*.

(A) Structure of the Cas9 constructs and mutagenesis efficiency. The number of transformants obtained for each construct is shown (Transf. nb). For the constructs with intronized Cas9, which produced many plants with mutant phenotypes, transformants from a part of the trays were uprooted for phenotype evaluation (70–109 randomly selected transformants). pNos, nopaline synthase promoter; Bar, coding sequence of the Bar gene; Ocst, octopine synthase terminator; Rps5a, *Arabidopsis* ribosomal protein 5a promoter; hCas9, human codon-optimized Cas9 coding sequence; zCas9, *Z. mays* codon-optimized Cas9 coding sequence; zCas9i, the same sequence with 13 introns represented as six yellow boxes; zCas9io, a sequence variant of zCas9; F, FLAG tag; N, nuclear localization signal; Nost, nopaline synthase terminator; U6p, *Arabidopsis* U6 promoter; T/C, target sequence of the guide RNA for *TRY* and *CPC* genes; gRNA, the conserved region of the guide RNA; LB and RB, left and right T-DNA borders.

(B) Picture of the primary transformants of two non-intronized Cas9 constructs, pAGM51523 and pAGM51535, and an intronized Cas9 construct pAGM51559. White arrows point to chimeric plants with leaves sectors with mutant phenotype.

Constructs of the *Z. mays* codon-optimized Cas9 containing two NLSs, pAGM51559 and pAGM51561, led to 72% (79/109) and 70% (59/84), respectively, of transformants with a full knockout phenotype. The intron-containing Cas9 with a single NLS produced 58% (41/70) of plants with a knockout phenotype.

Analysis of the transformants

To understand why some but not all plants displayed a mutant phenotype, DNA was extracted from primary pAGM51559 and

pAGM51561 transformants with a WT phenotype (eight plants for each construct). In addition, for both constructs, DNA was extracted from two plants with a full knockout phenotype and four plants with a chimeric phenotype. For each chimeric plant, two DNA extractions were made, one from leaves with a WT phenotype and one from leaves with a mutant phenotype. All DNA preparations were analyzed by PCR with two pairs of primers in the Cas9 construct, one for the amplification of the guide RNA region (primers guia1 and 4, product of 583 bp) and one for the amplification of a region within the Cas9 sequence

(primers casan2 and 3, product of 966 bp). Of the eight pAGM51559 transformants with the WT phenotype, one did not contain the Cas9 gene and guide RNA; we therefore hypothesized that it contained a truncated T-DNA (Supplemental Figure 4). The other seven plants with a WT phenotype appeared to contain a full T-DNA. For the eight pAGM51561 transformants, one carried an incomplete T-DNA containing the Cas9 gene but lacking the guide RNA region, while the other seven plants appeared to contain a full T-DNA. All other plants, either mutant or chimeric, contained both Cas9 sequences and the guide RNA.

DNAs prepared from the pAGM51559 and pAGM51561 transformants were analyzed by PCR with two pairs of primers designed to amplify the fragments of *CPC* and *TRY*, including the target site for Cas9. The PCR products were then sequenced (Supplemental Figures 5–8).

As expected, for pAGM51559 transformants, sequencing of the plant with the WT phenotype that lacked a complete T-DNA detected no mutation. Interestingly, two of the seven remaining pAGM51559 transgenic plants with a WT phenotype displayed some mutations in both *CPC* and *TRY* genes at the DNA level. These appeared as double peaks in the sequence traces starting at the Cas9 target site (Supplemental Figure 5A and 5B). The remaining five plants with a WT phenotype did not have any mutations at the DNA level. Analysis of the plants with a chimeric phenotype showed that all DNA samples, from leaves showing both WT and mutant phenotypes, displayed some mutations. Finally, as expected, the two transformants with a full knockout phenotype showed mutations at the DNA level in both *CPC* and *TRY* genes. To analyze the mutation spectrum in more detail, the PCR products of the *CPC* gene from five pAGM51559 transformants were cloned and sequenced. The PCR products selected for sequencing were from one transformant with the WT phenotype but contained mutations as previously determined by sequencing (named WT2), one transformant with a knockout phenotype (Mut9), and three transformants with a chimeric phenotype (DNA was extracted from leaves with a WT phenotype, CWT11, CWT12, and CWT14, Supplemental Figure 6). Sequences of the PCR products from the plant with a WT phenotype, WT2, were a mix of mutant and WT sequences. The same was observed for sequences of the WT leaves from chimeric plants. Finally, and as expected, analysis of sequences from plants with a full knockout phenotype showed the presence of only mutant sequences. In conclusion, plants displaying the WT phenotype in the entire plant or some sectors, both contained WT sequences, which explains their non-mutant phenotypes; however, some of these plants also contained mutations.

The same analysis performed on pAGM51559 transformants was also carried out in pAGM51561 transformants (Supplemental Figures 7 and 8). Of the seven plants that carried a complete T-DNA and exhibited a WT phenotype, four had mutations at the sequence level. Cloning and sequencing of the PCR products of five pAGM51561 transformants led to results and conclusions similar to those for pAGM51559 transformants. The overall spectrum of mutations included insertions of 1 bp at the cleavage site or deletions of 1, 3, or 7 bp.

In summary, constructs pAGM51559 and pAGM51561 that contained two NLSs and 13 introns yielded 70% and 72% of primary transformants with a full knockout phenotype, respectively, and 13% of transformants with a chimeric phenotype. In addition, one-quarter to one-half of the transformants with a WT phenotype displayed mutations at the DNA level. This means that 89% of the pAGM51559 transformants (79 + 15 + 4 estimated plants of 109) and 92% of the pAGM51561 transformants (59 + 11 + 7 estimated plants of 84) had an active Cas9 gene that was capable of generating mutations in the primary transformants.

Accumulation and localization of Cas9 with one or two NLSs in *Arabidopsis* and *N. benthamiana*

To understand why constructs with introns work better than those without introns and why two NLSs seem better than one, constructs similar to those described in the first experiment (Figure 1) but with Cas9 driven by the 35S promoter were used for transient expression in *N. benthamiana* leaves by *Agrobacterium*-mediated delivery. Cas9 protein accumulation was then analyzed by western blotting using a Cas9-specific antibody (Figure 2A, Supplemental Figure 9). Interestingly, in all cases, the amount of Cas9 protein with two NLSs was lower than that with one NLS expressed from corresponding genes (with the same nucleotide sequence, e.g., hCas9(1×NLS) versus hCas9(2×NLS)). We compared the accumulation of Cas9 with a single NLS, and it was similar between the expression from genes with different codon usage; however, protein amounts were slightly enhanced for the intron-optimized zCas9i gene (Figure 2A). This observation was further supported by results obtained for Cas9 with two NLSs: protein amounts evaluated by immunodetection clearly increased upon expression from intron-optimized genes, and there were no marked differences between zCas9i and zCas9io (Figure 2A).

The effect of intron optimization on protein accumulation was further tested by the immunodetection of Cas9 in primary *Arabidopsis* transformants using protein extracts prepared from either the pooled tissues of eight primary transformants (Figure 2B) or individual transformants (Figure 2C). For the assembly of pools, six plants with the WT phenotype and two plants with the chimeric phenotype were selected from transformations with constructs not containing introns, while eight plants with the mutant phenotype were selected for zCas9i (with one or two NLSs) and zCas9io. Overall, Cas9 levels appeared low in *Arabidopsis* but the immunodetection of Cas9 from pooled samples supported our previous observations with transient expression in *N. benthamiana*: lower protein levels were detected for Cas9 in fusion with two NLSs compared to Cas9 with a single NLS, and protein accumulation was enhanced upon expression from intron-containing genes (Figure 2B). Similarly, we failed to detect Cas9 in protein extracts from individual transformants carrying zCas9(2×NLS) without introns, but the protein was detected in three of the four plants transformed with the intron-optimized zCas9i with two NLSs (Figure 2C). Taken together, these results suggest that the addition of introns to the Cas9 gene leads to a higher expression level of the protein in *N. benthamiana* and *Arabidopsis*, or potentially more sustained expression.

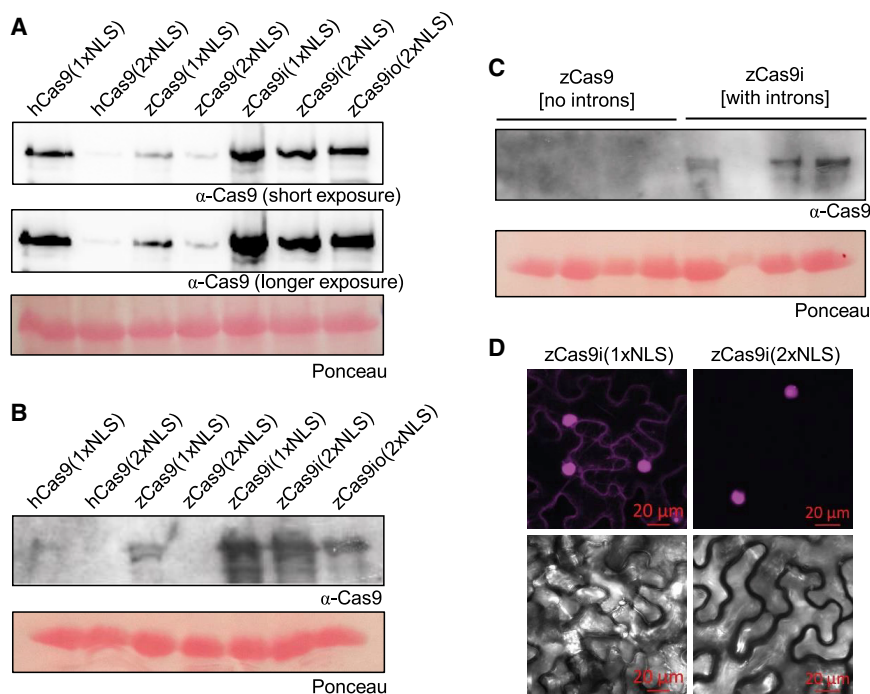


Figure 2. Cas9 protein accumulation observed after transient expression in *N. benthamiana* leaves and stable transformation in *Arabidopsis* plants.

(A) Accumulation of Cas9 upon the expression of different genes in *N. benthamiana*. Indicated Cas9 versions were expressed (under p35S control) in *N. benthamiana* by agroinfiltration. Tissues were used for protein extraction and immunodetection at 3 dpi. Ponceau staining of the membrane is shown as loading control.

(B) Accumulation of Cas9 upon expression from different genes in stable *Arabidopsis* transformants. Proteins were extracted from the pools of leaf tissues from eight independent primary transformants expressing each Cas9 version (with sgRNAs targeting *try* and *cpc*; 5-week-old plants) and used for SDS-PAGE and immunodetection. Pools of the first four samples (Cas9 without introns) were prepared from six WT-like and two chimeric plants. Pools of transformants expressing intron-optimized Cas9 variants were prepared from plants with mutant phenotypes (*try cpc*; hairy). Ponceau staining of the membrane is shown as the loading control.

(C) Detection of Cas9 in individual T₁ *Arabidopsis*

transformants. Proteins were extracted from the leaf tissues of individual T₁ transformants expressing either zCas9 or the intron-optimized zCas9i, both contained N- and C-terminal NLSs and were used for SDS-PAGE and immunodetection. Tissues exhibiting a WT-like phenotype were used for zCas9, and those with a *try cpc*-like phenotype were used for zCas9i. Ponceau staining of the membrane is shown as the loading control.

(D) Subcellular localization of Cas9 versions carrying one or two NLSs. As in **(A)**, except that fusions of Cas9 with mEGFP were expressed in *N. benthamiana*, tissues harvested 3 dpi were used for live cell imaging. The intron-optimized zCas9i gene was used for the expression of GFP-Cas9^{NLS} and NLS^{NLS}GFP-Cas9^{NLS} (NLS from SV40).

Although Cas9 with a single NLS accumulated to higher levels in *planta* compared to Cas9 with two NLSs, higher mutagenic activities were observed for the latter version (Figure 1). To analyze the subcellular localization of Cas9, two constructs coding for GFP-Cas9 fusions (GFP-zCas9^{NLS} and NLS^{NLS}GFP-zCas9^{NLS}) under the control of the 35S promoter were assembled and used for transient expression in *N. benthamiana* leaves. The GFP-Cas9 fusion protein containing a single NLS was observed in the nucleus as well as in the cytosol (Figure 2D). By contrast, the version with two NLSs was observed predominantly in the nucleus. This suggests that the mutagenic activity of Cas9 with a single NLS is limited by inefficient nuclear import. At the same time, the presence of two NLSs leads to a lower steady-state amount of the Cas9 protein, possibly due to enhanced turnover in the nuclear compartment. Intron optimization of the Cas9 gene appears to counteract this effect by enhancing overall expression, thus ensuring a nuclease concentration sufficient for efficient mutagenesis directly at the chromatin.

Test of a construct with low copy number in *Agrobacterium*

It is useful to be able to recover mutant plants that have lost the Cas9 construct in subsequent generations by segregation. If the T-DNA is present in one copy, one-quarter of the mutant plants in the next generation will lose the Cas9 construct. Obtaining plants lacking Cas9 becomes more difficult if several copies of the T-DNA are inserted into different chromosomal locations.

It is known that binary vectors that replicate at a single copy in *Agrobacterium* lead to a higher proportion of transformants containing single-copy, backbone-free transgenic plants (Ye et al., 2011). The plasmid backbone used for the experiments described above (Figure 1) has a pVS1 origin of replication, which is known to replicate in *Agrobacterium* at approximately 20 copies (Itoh et al., 1984; Heeb et al., 2000; Zhi et al., 2015). We therefore tested a binary vector (pAGM37443) that contains the *Agrobacterium rhizogenes* A4 plasmid origin of replication, which replicates in *Agrobacterium* at approximately one copy per cell (Nishiguchi et al., 1987; Ye et al., 2011). The same constructs as described previously (Figure 1) were assembled again in pAGM37443 and transformed into *Arabidopsis* by floral dipping. The number of transformants obtained was smaller than in the previous experiment (12–45 transformants for different constructs [Figure 3], compared with >80 transformants [Figure 1]). Lower transformation efficiencies were caused by seasonal effects and/or plant growth rather than differences in the copy number of constructs, as >170 primary transformants were obtained in an independent transformation experiment of one of the low-copy constructs, pAGM53451, which was repeated later (see below). Nevertheless, the same trend as before was observed: constructs with Cas9 without introns did not produce a full knockout phenotype in primary transformants; however, constructs with intronized Cas9 led to many primary transformants displaying a full knockout phenotype (47.6% and 50% for constructs with intronized Cas9 and two NLSs,

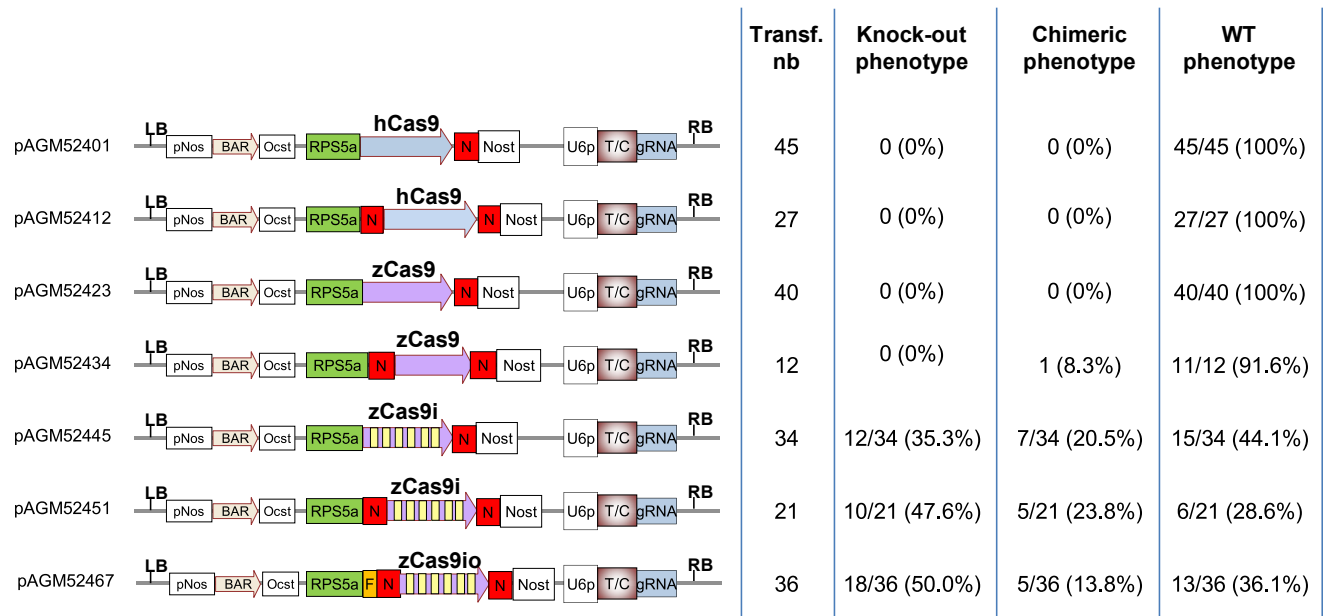


Figure 3. Comparison of different Cas9 versions in low-copy vectors by the mutagenesis of *TRY* and *CPC*.

Structure of the Cas9 constructs and mutagenesis efficiency. The legend for the annotations is the same as in Figure 1. pNos, nopaline synthase promoter; Bar, Bar gene coding sequence; Ocst, octopine synthase terminator; Rps5a, *Arabidopsis* ribosomal protein 5a promoter; hCas9, human codon-optimized Cas9 coding sequence; zCas9, *Z. mays* codon-optimized Cas9 coding sequence; zCas9i, zCas9 sequence with 13 introns represented as six yellow boxes; zCas9io, sequence variant of zCas9i; F, FLAG tag; Nost, nopaline synthase terminator; U6p, *Arabidopsis* U6 promoter; T/C, target sequence of the guide RNA for the *TRY* and *CPC* genes; gRNA, the conserved region of the guide RNA; LB and RB, left and right T-DNA borders.

respectively, Figure 3). The ratio of plants with a full knockout phenotype was slightly lower when using the *A. rhizogenes* A4 ori (47.6% and 50%) for intron-containing constructs and two NLS, versus 72% and 70% for -the same constructs with the pVS1 ori. This lower number of knockout plants may be explained by a lower number of plants containing multiple copies of the T-DNA integrated into the genome, as expected for plasmids that replicate at a single copy in *Agrobacterium*.

Because the first round of transformation did not yield many transformants with single-copy vectors, *Arabidopsis* transformation was performed again with construct pAGM53451. This time, 174 transformants were obtained, with 53 plants (30.4%) showing a WT phenotype, 50 plants (28.7%) exhibiting a chimeric mutant phenotype, and 71 plants (40.8%) displaying a full mutant phenotype. Therefore, constructs replicating at a single copy in *Agrobacterium* can produce a high number of transformants in *Arabidopsis* and still produce a high proportion of plants with a chimeric or full knockout phenotype.

Test of the intronized Cas9 for the mutagenesis of genes involved in flower development

To check whether significant edits can be obtained in other genes, guide RNAs were prepared for three genes involved in flower development: *APETALA3* (*AP3*; AT3G54340), *AGAMOUS* (*AG*; AT4G18960), and *LEAFY* (*LFY*; AT5G61850) (Yanofsky et al., 1990; Jack et al., 1992; Weigel et al., 1992). To facilitate the assembly of constructs, four cloning vectors that contain all components of the final constructs but not the target sequence of the guide RNA were generated (Figure 4A, Supplemental

Figure 10). Two of these vectors contain a BAR cassette for the selection of transgenic plants with Basta, and two contain a kanamycin-resistance cassette. The vector backbone contains either the pVS1 ori or the *A. rhizogenes* A4 ori. The missing sequence of the guide RNA cbe cloned easily by ligating two 24-nucleotide oligonucleotides into the *Bsa*I-digested vector. For the experiment described here, vector pAGM55261 (pVS1 ori) was used (Figure 4A, Supplemental Figure 10). The constructs for targeting *AP3* (pAGM55361), *LFY* (pAGM55961), and *AG* (pAGM55973) were transformed into *Arabidopsis*. We obtained 86 to 130 transformants for the three constructs. Before the plants flowered, 12 randomly selected transformants were transferred to single pots. All three transformations gave rise to a large number of plants with a knockout phenotype. All 12 transformants for all three target genes transferred to single pots gave rise to flowering mutant phenotypes. For *AP3*, all 12 plants had a complete knockout phenotype. We calculated non-transferred plants in the original tray (that were growing under crowded conditions) and identified some plants with the WT phenotype (3 of more than 32 transformants flowering at the time of counting). For *LFY*, 10 of the 12 plants showed a full knockout phenotype and 2 were chimeric, with a large part of the plants displaying the mutant phenotype and only a few branches showing a WT phenotype, which allowed the plants to set seeds. For *AG*, 11 plants had a knockout phenotype and 1 displayed a chimeric phenotype. In summary, full mutant phenotypes were observed with a frequency greater than 80% when targeting *AP3*, *AG*, or *LFY* with independent constructs, each containing one guide RNA and zCas9i. Thus, intron-optimized Cas9 appears to perform robustly at multiple different loci and with different guide RNAs.

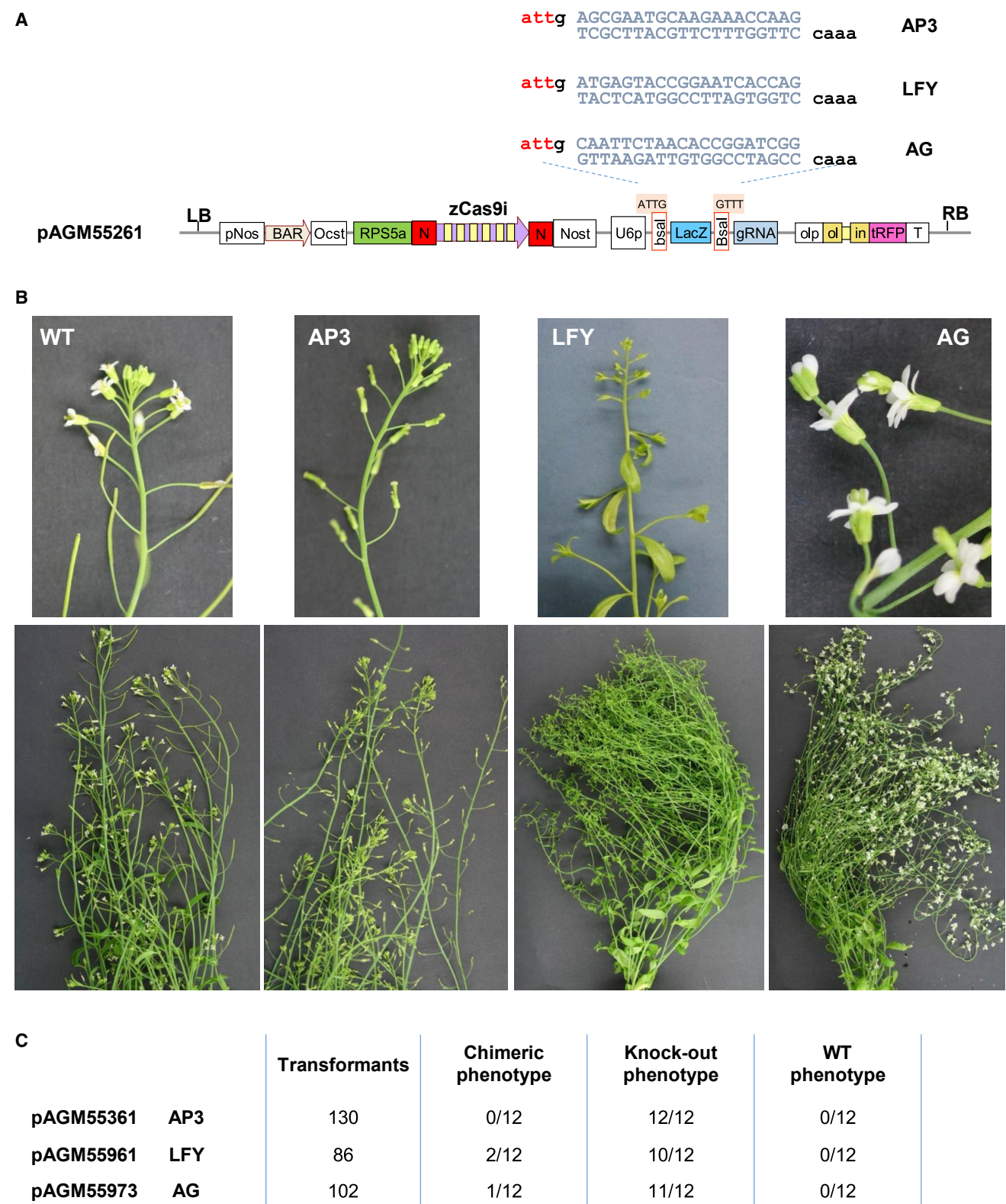


Figure 4. Mutagenesis of floral homeotic genes using intronized Cas9 constructs.
(A) Schematic representation of the constructs and cloning strategy. For each construct, two oligonucleotides were ligated into the cloning vector pAGM55261 digested with *Bsa*I. The resulting constructs were transformed into *Arabidopsis* by the floral dip method.
(B) Phenotypes of the wild-type control and transformants.
(C) Estimation of the number of transformants with wild-type and mutant phenotypes.

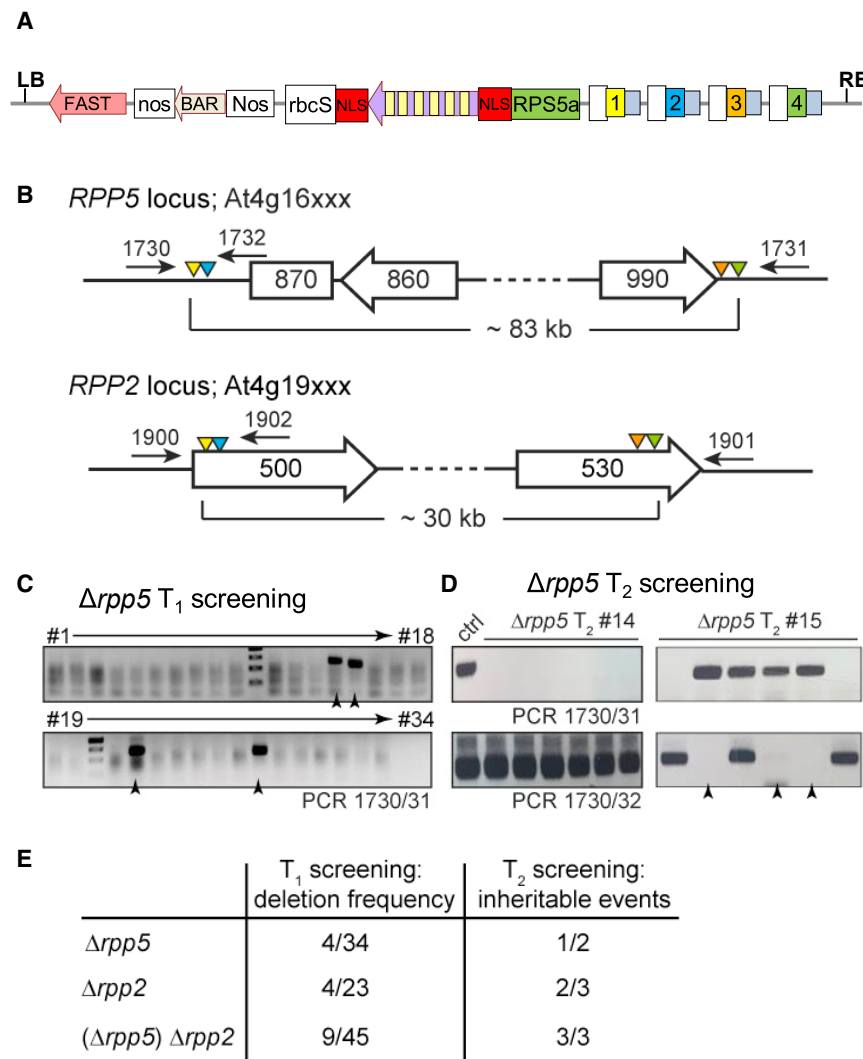


Figure 5. Generation of large chromosomal deletions in *Arabidopsis* using the intron-optimized zCas9i.

(A) General architecture of the constructs used for generating large deletions. sgRNAs (under the control of pU6-26) are transcribed in the opposite direction relative to zCas9i. FAST, pOle1:Ole1-tagRFP_toe1 (Shimada et al., 2010).

(B) Schematic drawing of the *RPP5* and *RPP2* loci with sgRNA target sites (triangles; color code corresponds to [A]) and primer binding sites indicated. Genes are indicated by arrows with *Arabidopsis* gene identifiers, and a transposable element is shown as a rectangle. Not drawn to scale.

(C) Screening of T₁ plants for candidate lines carrying a deletion at the *RPP5* locus. T₁ plants were selected by resistance to BASTA and screened by PCR with the indicated primers for the occurrence of a deletion at the *RPP5* locus. Thirty-four independent T₁ lines were screened and putative deletion lines are marked with arrowheads. A PCR product of ~590–750 bp was expected for deletion lines.

(D) PCR screening of segregants from T₂ families 14 and 15 shown in (C) for the isolation of $\Delta rpp5$ deletion line. $\Delta rpp5$ deletion alleles detected by PCR are shown at the top (oligonucleotides 1730/31; product size 592 bp; see also Supplemental Figure 11). The PCR product from a T₁ plant served as the control (ctrl). Wild-type *RPP5* locus detected by PCR is shown at the bottom (1730/32; product size 606 bp). PCR products of Col DNA served as the control.

(E) Summary of frequencies of large deletions in *Arabidopsis* generated by intron-optimized zCas9.

Generation of large chromosomal deletions in *Arabidopsis* with intron-optimized Cas9

We previously generated large chromosomal deletions ranging from 70 to 120 kb using intron-less hCas9 driven by a ubiquitin promoter (Ordon et al., 2017, 2019). In these experiments, deletions occurred at low frequencies (0.5%–2%) and could only be detected in the T₂ generation. The efficiency of zCas9i (under *pRPS5a* control) for the generation of chromosomal deletions was tested by targeting *RPP5* and *RPP2* gene clusters (Figure 5B). Two different constructs, each containing four guide RNAs, were assembled (Figure 5A) to program Cas9 for two target sequences on each side of the respective cluster (Figure 5B). Constructs were similar to those used in previous experiments (Figure 1), except that they contained the FAST marker (Shimada et al., 2010) for transgene counterselection by seed fluorescence. The frequency of deletion events was evaluated directly in primary T₁ transformants by PCR screening (Figure 5C). Deletions with an expected size of ~83 kb were detected with a frequency of ~10% (4/34) at the *RPP5* locus (Figure 5C). The *RPP2* locus (~30 kb) was targeted for deletion both in WT Col plants and in

the background of a $\Delta rpp5$ mutant line. The expected deletion was detected in 15%–20% of the screened primary transformants (Figure 5E). Furthermore, the inheritability of deletions detected in T₁ was tested in transgene-free T₂ segregants, which were selected by the absence of seed fluorescence (Figure 5D). Chromosomal deletions detected in the T₁ generation were inheritable for six of the eight families (Figure 5E). Therefore, these deletions likely occurred early in development, as also observed phenotypically when targeting *try cpc*, and were thus efficiently transmitted to the germline. The deletion of the targeted gene clusters was further confirmed by sequencing the junction sites in selected mutant lines, and *rpp5* and *rpp2* deletion lines lost resistance to *Hyaloperonospora arabidopsidis* isolates Emwa1 (resistance conferred by *RPP4* within the *RPP5* cluster in accession Col; van der Biezen et al., 2002), and Cala2 (resistance conferred by *RPP2*; Sinapidou et al., 2004), as expected (Supplemental Figure 11). Taken together, intron-optimized Cas9 induces chromosomal deletions at relatively high frequencies (10%–20% for deletions ranging up to 80 kb), and deletion lines can be easily isolated from the T₂ segregants of as few as two or three PCR-confirmed T₁ plants.

Test of the intronized Cas9 in *N. benthamiana*

The human-codon-optimized Cas9 and the intronized zCas9i with two NLSs were further tested in *N. benthamiana*. Four constructs containing either hCas9 or zCas9i together with guide RNA rt1 or rt4 were designed to target two *N. benthamiana* genes with putative rhamnosyltransferase activity (Supplemental Figure 12). With guide RNA rt1, all primary transformants obtained (hCas9 4/4, zCas9i 11/11) had mutations at the target site (analyzed by the amplification of target-site sequences and the sequencing of the PCR product). This suggests that both Cas9 genes can mediate genome editing with high efficiencies in *N. benthamiana*. This result is in agreement with previous reports on efficient editing in this species with hCas9 (Ordon et al., 2017; Castel et al., 2018; Adachi et al., 2019). By contrast, no mutations were detected in any of the putative transformants (hCas9 0/7, zCas9i 0/6) when using rt4 (targeting a different gene), suggesting that the guide RNA for this selected target site did not work.

Eight additional target sites in either rhamnosyltransferase genes or a flavonol synthase gene were tested with the intronized Cas9 only (Supplemental Figure 12). For one of these (guide fls3), no mutation was obtained in 14 putative transformants with the intronized Cas9 gene. For the other seven targets, mutations were obtained in some plants, with frequencies ranging from 25% to 100% for primary transformants with mutations (Supplemental Figure 12). However, these numbers are not fully quantitative, as we did not test whether putative transformants without mutations in target sites were true transformants or contained complete T-DNAs.

In addition, the induction of small deletions by targeting multiple sites within a target gene was attempted in *N. benthamiana* using zCas9i with two NLSs. In these experiments, a binary vector with the pVS1 origin of replication was used. Three different genes, *Roq1* (Recognition of XopQ 1), *NRG1* (N requirement gene 1), and an *NPR1* (Nonexpressor of Pathogenesis-Related Genes 1) homolog, were targeted by multiplexing with two (*NRG1*) or three (*Roq1*, *NPR1*) different sgRNAs (Figure 6A; see Methods for target sites). The distance between the outmost expected cut sites in the targeted genes varied from 57 nt (*NRG1*) to 265 nt (*Roq1*). DNA was extracted from primary transformants and employed for PCR screening using primers flanking the target sites (Figure 6B). All three groups of transformants contained several plants with PCR-detectable deletions between the target sites (*Roq1* 7/10, *NRG1* 7/10, *NPR1* 6/11). Notably, a PCR product corresponding to the WT size was no longer detectable in several plants due to the editing of *Roq1* and the *NPR1* homolog (Figure 6B; red arrows). More than two bands were amplified from several transformants (Figure 6B; blue arrows), suggesting somatic chimerism of the tissues used for DNA extraction. Taken together, the data from the generation of point mutations and deletions indicates that the intronized Cas9 mediates the efficient induction of double-strand breaks at target sites in *N. benthamiana*.

Gene knockout in the hairy roots of *C. roseus*

The intronized Cas9 was also tested in another plant species, the Madagascar periwinkle, *C. roseus*, to mutate and delete three genes, the Jasmonate-Associated MYC2 (JAM) tran-

scription factors *JAM2* and *JAM3*, as well as Repressor of MYC2 Targets 1, *RMT1* (Figures 7, Supplemental Figure 13). Two constructs were prepared using the intronized Cas9 with two NLSs (zCasi), which was under the control of an *Arabidopsis* ubiquitin promoter (*AtUBQ10*) in a vector with a pVS1 ori. The first construct, pSB311, contained eight guide RNAs (under the control of the pU6-26 promoter) targeting two genes (*JAM2* and *JAM3*), while the second construct, pSB312, contained eight guide RNAs targeting three genes (*JAM2*, *JAM3*, and *RMT1*). The third construct, pSB310, served as the control. It expressed the intronized Cas9 (zCasi) not containing sgRNAs. After *Agrobacterium*-mediated transformation and selection with hygromycin B, independent transgenic hairy root lines were generated (30 lines with pSB311, 19 lines with pSB312, 26 lines with pSB310), and the best-growing lines were transferred from plates to liquid culture media. A subset of these hairy root lines was tested for transgene integration: 19/19 were positive for the hygromycin B resistance gene and 14/19 had the guide RNA region.

Of the confirmed lines, four lines for each construct (pSB311 and pSB312) were further characterized for mutations in the genes of interest. If different-sized PCR products were obtained for a line, these PCR products were directly sequenced. If only one PCR product was obtained, it was sequenced to test whether different mutations were present. If different mutations were identified, the products were cloned to separate the mutations, and individual colonies were sequenced to identify each mutation. In total, eight independent lines were tested for two or three alleles at eight target sites. We observed 100% editing at the target sites (Figure 7, Supplemental Figure 13). Some lines seemed to be homozygous, suggesting that one of the two parental genes may have been repaired after the cleavage of Cas9 by gene conversion using the other mutated homolog as template. One line appeared chimeric with three different alleles. This occurs when both homologs are not cleaved and repaired immediately in the first cell after transformation but after the division of the initially transformed cell. The spectrum of mutations observed included the insertion of 1 nucleotide, deletions of 1 to 64 nucleotides, and the deletion of all sequences between some pairs of sgRNAs. Two control lines that contained the Cas9 construct but not guide RNAs were also sequenced for the region of the gene of interest and no mutations were found. Thus, zCas9i is also highly efficient for multiplex editing in *C. roseus* hairy roots, with an apparent mutagenesis efficiency of 100%.

DISCUSSION

We show here that the codon composition of the Cas9 sequence, the number of encoded NLSs, and the presence of introns in the coding sequence can affect the expression level of the Cas9 nuclease and the final DNA cleavage efficiency in transformants. However, the effect of codon optimization on Cas9 nuclease activity in our study was minor. Two Cas9 genes with either human or maize codon-optimized sequences were compared (Figure 1A). Both versions (lacking introns) displayed very low activity, with no full knockout phenotypes observed in primary transformants and only a few of the transformants showing mutant leaf sectors. Using Cas9 versions containing two NLSs,

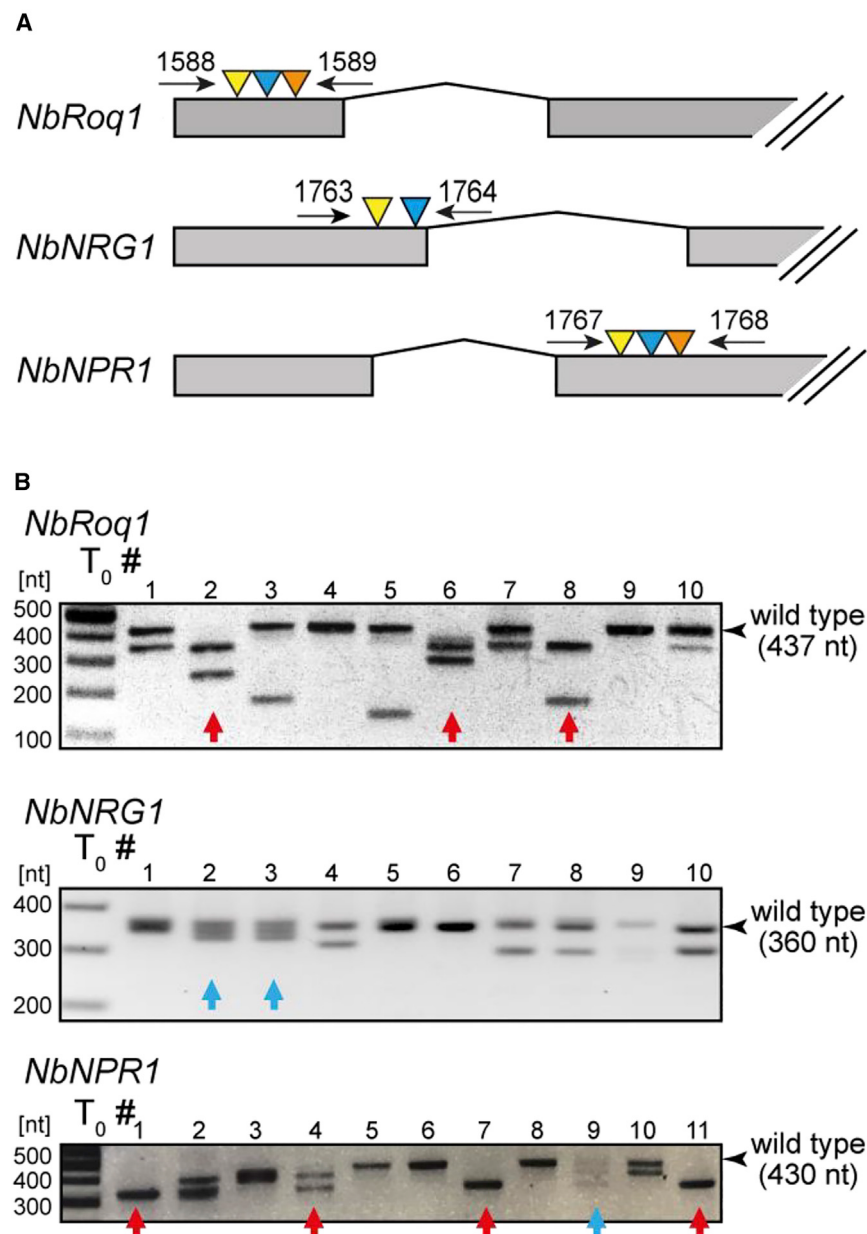


Figure 6. Frequency of small deletions in *N. benthamiana* generated by zCas9i.

(A) Gene models of loci targeted for the induction of deletions by multiplex genome editing in *N. benthamiana*. Constructs used for editing the depicted loci had architectures similar to that shown in Figure 5A, except that zCas9i expression was driven by the 35S promoter. *Roq1*, GenBank entry MF773579.1 (Schultink et al., 2017); *NRG1*, Niben101Scf02118g00018, GenBank entry DQ054580.1; *NPR1*, NPR1-like gene, Niben101Scf14780g01001.

(B) Genotyping of primary (T_0) transformants by PCR. For each transformation, results from 10 or 11 randomly chosen T_0 individuals are shown, and genotyping was conducted using oligonucleotides indicated in (A). Red arrows mark individuals lacking a PCR product corresponding to the wild-type fragment. Blue arrows mark individuals with multiple (>2) PCR products, which indicate somatic chimerism.

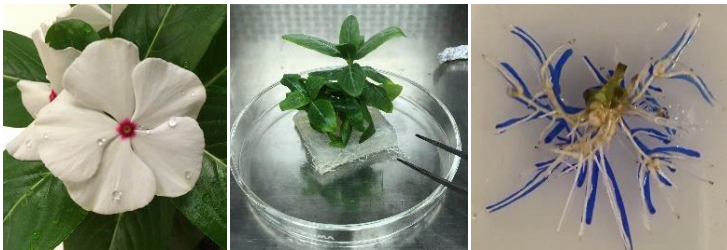
(Koblan et al., 2018; Maggio et al., 2020). In plants, many versions of Cas9 with either one or two NLSs have been reported to be capable of genome editing (Belhaj et al., 2013), but the importance of nuclear import was not characterized in detail. In this work, constructs with the same Cas9 sequence containing either one or two NLSs were systematically compared to evaluate the impact of the presence of one or two NLSs. When using Cas9 genes without introns, the efficiency for obtaining plants with a chimeric phenotype increased slightly from 1% with one NLS to 3.2% with two NLSs (Figure 1A). In addition, a construct with the intron-containing maize-codon-optimized sequence with two NLSs (pAGM51559) yielded 72% plants with a knockout phenotype in primary transformants compared with only 58% when using a similar construct containing only one NLS (pAGM51547, Figure 1). The same trend was also observed with similar constructs in low-

copy vectors (pAGM52541 and pAGM52445, Figure 3). Interestingly, the transient expression of the Cas9 gene in *N. benthamiana* and the analysis of Cas9 accumulation in primary *Arabidopsis* transformants indicated that the steady-state amount of Cas9 in cells was lower for the version that had two NLSs (Figures 2A and 2B). This suggests that the turnover rate of Cas9 in the nucleus may be higher than that in the cytoplasm. It is known that the turnover of some proteins that contain degradation signals can be higher in the nucleus than in the cytosol (Lenk and Sommer, 2000). In yeast, this was linked to the higher ubiquitination rate upon the nuclear import of the protein. Although Cas9 is not naturally present in eukaryotes, a similar phenomenon may be taking place when the Cas9 protein is transported to the nucleus. We also demonstrated that a GFP-Cas9 fusion containing two NLSs was predominantly observed in the nucleus, while that with a single NLS was observed in both

such chimeric plants were obtained in slightly higher frequencies with the maize-codon-optimized version than with the human-codon-optimized version (Figure 1). A small effect of codon optimization on Cas9 activity was also reported in a study comparing the same human-codon-optimized version that we used in this work with an *Arabidopsis*-codon-optimized Cas9 (Castel et al., 2019). However, it is possible that in some cases, codon optimization alone may be sufficient to increase the efficiency in target organisms, as codon-optimized Cas9 genes have been used successfully in several species, such as rice (Zhou et al., 2014; Mikami et al., 2015).

Nuclear import of the Cas9 nuclease is critical for the cleavage of chromosomal DNA in eukaryotic cells. In animal systems, a substantial improvement in Cas9 nuclear localization and activity has been reported for Cas9 versions containing up to four NLSs

A



B

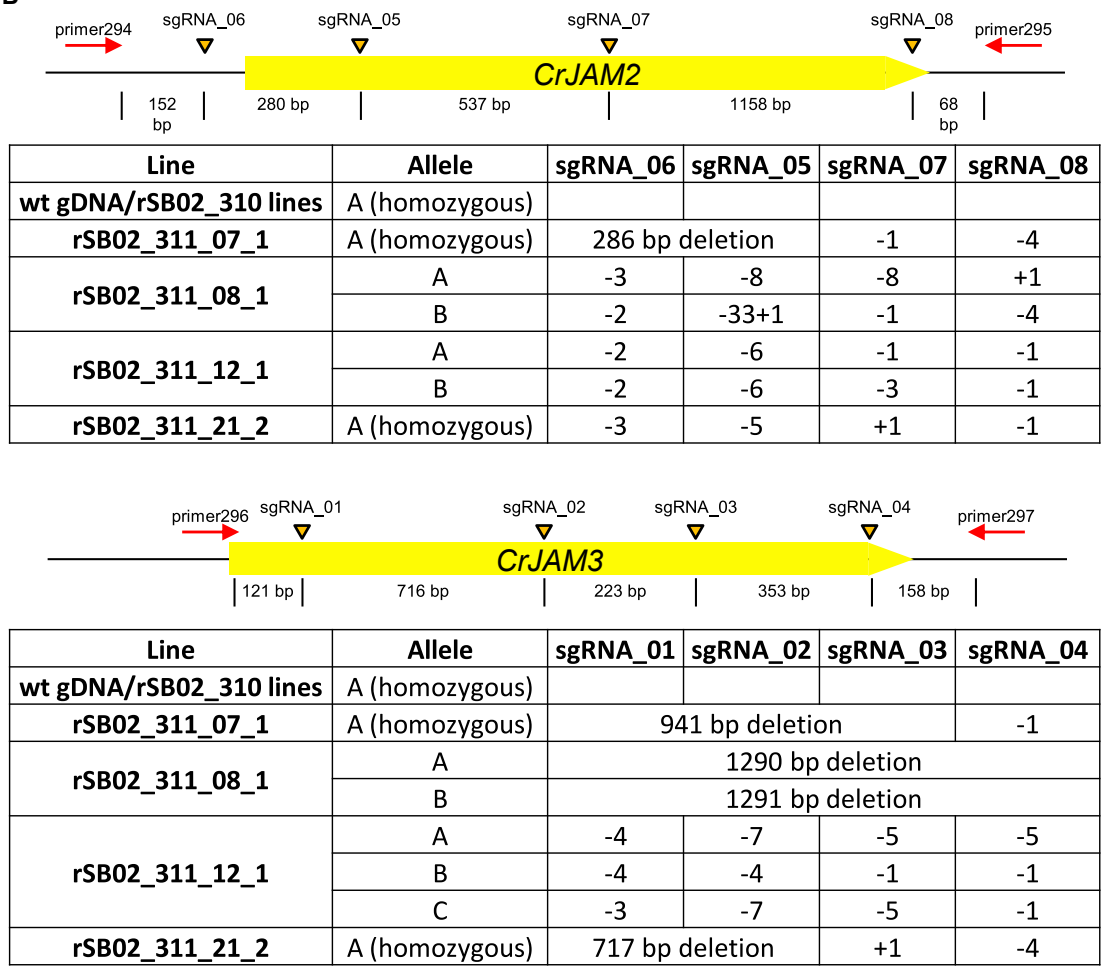


Figure 7. Site-targeted mutagenesis in *C. roseus* using the intronized Cas9.

(A) Picture of a wild-type plant with flower, an *in vitro*-grown plant used for transformation, and a nodal segment with emerging hairy roots. plant. (B) Summary of the mutations generated by multiplexing in transgenic hairy root lines of *CrJAM2* and *CrJAM3*. The genomic DNA of each line was extracted, the alleles were PCR-amplified and cloned into *Escherichia coli* (*E. coli*), and *E. coli* colonies were sequenced with the indicated primers. Most lines had two alleles with different mutations in both parental sequences (designated alleles A and B; *C. roseus* is a diploid species). In one line, a third allelic sequence was identified (allele C), probably showing two cell lineages with different mutations in one of the two homologs. In some cases, only one allele was detected (noted as homozygous). This may have been due to the repair of one of the two parental sequences by gene conversion using the other mutated homolog as the template.

the nucleus and cytosol (Figure 2D). This suggests that a single NLS at the C-terminal end is not sufficient for the efficient transfer of the Cas9 enzyme to the nucleus, thus limiting nuclease activity at the chromatin level.

The most important feature of the Cas9 gene for the efficient generation of site-targeted mutations, at least in *Arabidopsis*, was the presence of multiple introns in Cas9 coding sequences. This factor alone was sufficient to convert a weakly active Cas9

construct that produced no plants with a knockout phenotype in primary transformants into a very efficient construct that led to 70% to 100% of transformants with a mutant phenotype when targeting one or two genes with a sgRNA (Figures 1, 3, and 5; the trichome phenotype used to compare different Cas9 versions depends on the inactivation of all copies of the *TRY* and *CPC* genes—four alleles). There are two possible explanations for the improved editing efficiency that is obtained by the introduction of introns into the Cas9 gene in *Arabidopsis*. First, intron optimization leads to higher accumulation of the Cas9 nuclease, as shown for both transient expression in *N. benthamiana* and stable transformation in *Arabidopsis* (Figure 2). This effect may be caused by enhanced gene expression and/or nuclear export of the transcript (Shaul, 2017). In addition, the inclusion of introns into the Cas9 gene may counteract transgene silencing and lead to more sustained expression of the nuclease during development (Christie et al., 2011). Indeed, we rarely observed chimeric plants in previous experiments using mainly hCas9 (Ordon et al., 2017, 2019), but they regularly occurred with zCas9i. However, it should be noted that most phenotypes (e.g., the hairy phenotype of *try cpc* plants, Figure 1B) occurred early in development and were already visible in first emerging organs.

Although the addition of introns to the Cas9 gene transformed weakly active constructs into efficient ones, it is likely that efficient constructs can be made with Cas9 lacking introns by optimizing the Cas9 sequence alone. Indeed, some Cas9 genes that do not contain introns have been reported to generate high numbers of transgenic plants with mutant phenotypes (Belhaj et al., 2013; Tsutsui and Higashiyama, 2017). A test of different versions of the nucleases within otherwise identical constructs will help to make precise comparisons of these sequences.

The presence of introns in the Cas9 sequence may not be critical for efficient editing in all plant species. For example, the human-codon-optimized version we used here was reported to work well in some plants such as tomato (e.g., Brooks et al., 2014). Nonetheless, the enhanced Cas9 protein accumulation in *N. benthamiana* with the zCas9i gene suggests that positive effects from intron optimization are not limited to *Arabidopsis*, and high mutation rates were obtained in stable *N. benthamiana* lines (Figure 6) and *C. roseus* hairy roots (Figure 7). In conclusion, an efficient Cas9 construct can be generated by inserting multiple introns into the coding sequence, by optimizing the Cas9 gene sequence to achieve a high level of Cas9 expression in a given species, or both. Introns may not be necessary for achieving a high expression level but can lead to a spectacular improvement in the efficiency of the constructs for Cas9 sequences that do not work well enough on their own.

The intron-containing zCas9i performed extremely well in combination with the *RPS5a* promoter (e.g., Figure 1A). Surprisingly, we did not obtain genome-edited plants when using zCas9i in combination with the egg-cell-specific *EC1.2/DD45* or the *ECenH* promoter fragments and the pea E9 terminator as recommended by Wang et al. (Wang et al., 2015; Castel et al., 2019) (data not shown). In a similar manner, Cas9 driven by the egg-cell-specific promoter also did not yield any mutant lines when used in combination with the Nos terminator, which is a strong terminator in many cases (Wang et al., 2015). It is possible that in the

egg cell, specific combinations of promoters, Cas9 gene version, and terminators are needed to yield a very specific amount of Cas9 protein for activity, and it is possible that the egg-cell-specific promoter combined with the intronized Cas9 gene and yet another terminator would work perfectly well. The intronized Cas9 also worked well with the *Arabidopsis* ubiquitin promoter in *C. roseus*. To enable users to make constructs with the intronized Cas9 coding sequence with either the *RPS5a* promoter or other promoters, a series of modules and Cas9 vectors was deposited at Addgene (Supplemental Figures 10 and 15–18).

In summary, adding introns to a Cas9 gene can significantly improve the efficiency of Cas9 constructs for mutagenesis. Cas9 has also been used for other applications such as transcriptional activation or repression. One can speculate that adding introns to these Cas9 versions may have as strong an effect as for use as a nuclease, and potentially even more, as a sustained high level of expression is required for transcriptional repression.

METHODS

Plasmid construction

The Cas9 gene sequence was optimized for the codon usage of *Z. mays* using software from Eurofins MWG. To introduce introns into the Cas9 gene sequence, several *Arabidopsis* introns were randomly selected from genome sequences, and those that were well recognized by the NetGene2 online software (<http://www.cbs.dtu.dk/services/NetGene2/>) (Hebsgaard et al., 1996) when inserted into Cas9 sequences were selected. In some cases, splice-site sequences were modified to be closer to *Arabidopsis* splice-site consensus sequences (Brown et al., 1996). Most of the selected introns had previously been used successfully to improve a tobacco mosaic virus vector for use in *N. benthamiana* (Marillonnet et al., 2005). The addition of the 13 introns to the Cas9 gene adds 1.6 kb to the 4.1 kb Cas9 coding sequence, resulting in a total coding region of 5.7–5.8 kb, depending on the presence of one or two NLSs. This relatively small increase in size did not affect the cloning efficiency of the constructs. The Cas9 coding sequence was then constructed using a combination of gene synthesis with DNA fragments obtained from Eurofins MWG and PCR amplification from *Arabidopsis* genomic DNA for the introns. As the Cas9 sequence is quite large, six subparts with or without introns were cloned as level 1 modules using Golden Gate cloning and the MoClo system. The subparts were then assembled as level 0 modules, which contained the entire coding sequence with either one or two NLSs, yielding modules pAGM12591 and pAGM47539 (*Z. mays* codon-optimized Cas9 without introns and with one or two NLSs, respectively), as well as pAGM13741, pAGM47523, and pAGM51073 (*Z. mays* codon-optimized Cas9 with 13 introns and with one, two, and two NLSs, respectively, Supplemental Figure 1). pAGM47523 and pAGM51073 differ from pAGM51073 in an N-terminal FLAG tag and mutations at four sites in introns 1, 3, 12, and 13, which were designed to remove potential cryptic splice sites in nearby exon sequences (Supplemental Figure 1). The complete sequences of different Cas9 gene sequences are provided in Supplemental Information (Supplemental Figure 14). The final Cas9 expression constructs in Figures 1 and 3 were made using the MoClo system (Engler et al., 2014; Marillonnet and Grutzner, 2020).

Plasmids containing a Cas9 expression cassette, a transformation marker, and two *BsaI* sites for the cloning of sgRNA (pAGM55261, pAGM55273, pAGM55285, and pAGM55297, Supplemental Figure 10) are available from Addgene. In addition, vectors that already contain Cas9 (pAGM65879, pAGM65881, pAGM65893, and pAGM65905),

empty binary vectors (pAGM47443 and pAGM41991), the intronized Cas9 coding sequence as level 0 module (pAGM47523, zCas9i), and plant transformation selection markers as level 1 modules (pAGM35171 and pAGM67131) are also available from Addgene (list in [Supplemental Figure 10](#)). This will allow users to prepare constructs with other promoters and multiple guide RNAs using the MoClo system (<https://www.addgene.org/kits/marillonnet-moclo/>). The strategy used for assembling Cas9 constructs using the MoClo system is provided in [Supplemental Figures S15–S18](#).

Plant transformation, growth, and selection and infection assays

Arabidopsis Columbia-0 (Col) plants were transformed using *Agrobacterium* strains by the flower dip method ([Clough and Bent, 1998](#)). *N. benthamiana* plants were transformed as previously described ([Gantner et al., 2019](#)); a detailed protocol is provided online ([dx.doi.org/10.17504/protocols.io.sbaeiae](https://doi.org/10.17504/protocols.io.sbaeiae)). *C. roseus* hairy roots were generated as described in [Rizvi et al. \(Mortensen et al., 2019\)](#) with the modification made in [Mortensen et al. \(Mortensen et al., 2019\)](#). The infection of plants with *H. arabidopsidis* and trypan blue staining were conducted as previously described ([Stuttman et al., 2011](#)).

Guide RNA design

For the mutagenesis of *CPC* and *TRY*, a validated guide RNA target sequence was used ([Wang et al., 2015](#)): G aatatctctatctctc tgg (protospacer adjacent motif underlined). The first G is not in the target site and serves as the transcription start site for the *Arabidopsis* U6 promoter. Guide RNAs for genes regulating flower morphology (*AP3*, *LFY*, *AG*) and for the generation of large deletions (*RPP5* and *RPP2* clusters) were selected using CHOPCHOP (<https://chopchop.cbu.uib.no/>) ([Labun et al., 2019](#)). sgRNA target sites for editing in *N. benthamiana* were selected using CRISPR-P 2.0 ([Liu et al., 2017](#)). *C. roseus* sgRNA sequences were selected using Benchling ([benchling.com](#)), and on-target scores were predicted ([Doench et al., 2016](#)).

Primers used for genotyping the transformants

To detect the presence of T-DNAs in transformed plants, primer pairs casan2 (cagctgcagcttcgctcctc)/casan3 (ggcgttatcaccgatgaat acaagg) and guia1 (ccaaggaagtgttgacgctacc)/guia4 (cacttttcaagttga taacggactagcc) were used to amplify a Cas9 gene fragment and the guide RNA region, respectively. To analyze the *CPC* target locus, the *CPC* gene was amplified by PCR using primers cpcan1 (tt ggtctc a ACAT gtcagaact cactttggctagtgttg) and cpcan2 (tt ggtctc a ACAA atcatgtgtcgtggagg ctgg) and sequenced with primer cpcan1. To analyze the *TRY* target locus, the *TRY* region was amplified by PCR using primers tryan1 (tt ggtctc a ACAT ggggaagcacatgtgttcac) and tryan2 (tt ggtctc a ACAA gttgtggatat aaaagtcgtagacgag) and sequenced with primer tryan2.

Immunodetection and live cell imaging

Protein extracts were prepared by grinding the tissues in liquid nitrogen and boiling in Laemmli buffer. Total extracts were resolved on 6% or 8% SDS–PAGE gels, and the proteins were transferred to nitrocellulose membranes. A rabbit monoclonal α -Cas9 antibody (Abcam EPR18991) was used together with an HRP-conjugated secondary antibody (GE Healthcare), and SuperSignal West Pico and West Femto chemiluminescence substrates (Thermo Fisher) were used for detection. The subcellular localization of GFP–Cas9 variants was determined under a Zeiss LSM780 confocal laser scanning microscope. GFP was excited using the 488 nm laser, and the detection range was set to 493–532 nm.

SUPPLEMENTAL INFORMATION

Supplemental Information is available at *Plant Communications Online*.

FUNDING

The work of S. Marillonnet, R.G., and C.H. was supported by core funding of the IPB. Work in the JS lab was supported by grant STU642-1/1 (DFG/

GRC; Deutsche Forschungsgemeinschaft) to JS. Work on *C. roseus* was supported by the National Science Foundation MCB award 1516371 to C.L.-P. and E.C.

AUTHOR CONTRIBUTIONS

S. Marillonnet designed the intronized Cas9 gene and the constructs, as well as the experiments for comparing and testing Cas9 in *Arabidopsis* and *N. benthamiana*. J.S. designed and tested the Cas9–GFP fusion, supervised the Cas9 quantification experiments, and designed constructs and experiments for generating deletions in *Arabidopsis* and *N. benthamiana*. R.G., C.H., and P.M. performed experiments. S. Mortensen, E.J.C., and C.L.P. designed and S. Mortensen performed the *Catharanthus* experiments. S. Marillonnet wrote the manuscript with contributions from J.S. and all authors.

ACKNOWLEDGMENTS

We thank Nicola Patron and Jonathan D.G. Jones from the Sainsbury lab at the John Innes Centre in Norwich for providing the hCas9 gene module and the level 0 construct that contains a consensus sequence of the *Arabidopsis* U6 promoter (pICSL90001). J.S. acknowledges Carola Kretschmer and Bianca Rosinsky for excellent technical assistance and greenhouse work. The authors declare no conflicts of interest.

Received: July 1, 2020

Revised: November 13, 2020

Accepted: November 19, 2020

Published: November 23, 2020

REFERENCES

- Adachi, H., Contreras, M., Harant, A., Wu, C.-h., Derevnina, L., Sakai, T., Duggan, C., Moratto, E., Bozkurt, T.O., Maqbool, A., et al. (2019). An N-terminal motif in NLR immune receptors is functionally conserved across distantly related plant species. *bioRxiv*, 693291.
- Ahmad, N., Rahman, M.U., Mukhtar, Z., Zafar, Y., and Zhang, B. (2020). A critical look on CRISPR-based genome editing in plants. *J. Cell Physiol.* **235**:666–682.
- Belhaj, K., Chaparro-Garcia, A., Kamoun, S., and Nekrasov, V. (2013). Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods* **9**:39.
- Brooks, C., Nekrasov, V., Lippman, Z.B., and 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 Physiol.* **166**:1292–1297.
- Brown, J.W., Smith, P., and Simpson, C.G. (1996). *Arabidopsis* consensus intron sequences. *Plant Mol. Biol.* **32**:531–535.
- Callis, J., Fromm, M., and Walbot, V. (1987). Introns increase gene expression in cultured maize cells. *Genes Dev.* **1**:1183–1200.
- Carle-Uriste, J.C., Brendel, V., and Walbot, V. (1997). A combinatorial role for exon, intron and splice site sequences in splicing in maize. *Plant J.* **11**:1253–1263.
- Castel, B., Ngou, P.M., Cevik, V., Redkar, A., Kim, D.S., Yang, Y., Ding, P., and Jones, J.D.G. (2018). Diverse NLR immune receptors activate defence via the RPW8-NLR NRG1. *New Phytol.* **222**:966–980.
- Castel, B., Tomlinson, L., Locci, F., Yang, Y., and Jones, J.D.G. (2019). Optimization of T-DNA architecture for Cas9-mediated mutagenesis in *Arabidopsis*. *PLoS One* **14**:e0204778.
- Christie, M., Croft, L.J., and Carroll, B.J. (2011). Intron splicing suppresses RNA silencing in *Arabidopsis*. *Plant J.* **68**:159–167.
- Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**:735–743.

- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**:819–823.
- Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**:184–191.
- Engler, C., Youles, M., Gruetzner, R., Ehnert, T.M., Werner, S., Jones, J.D., Patron, N.J., and Marillonnet, S. (2014). A golden gate modular cloning toolbox for plants. *ACS Synth. Biol.* **3**:839–843.
- Gantner, J., Ordon, J., Kretschmer, C., Guerois, R., and Stuttmann, J. (2019). An EDS1-SAG101 complex is essential for TNL-mediated immunity in *Nicotiana benthamiana*. *Plant Cell* **31**:2456–2474.
- Hebsgaard, S.M., Korning, P.G., Tolstrup, N., Engelbrecht, J., Rouze, P., and Brunak, S. (1996). Splice site prediction in *Arabidopsis thaliana* pre-mRNA by combining local and global sequence information. *Nucleic Acids Res.* **24**:3439–3452.
- Heeb, S., Itoh, Y., Nishijyo, T., Schnider, U., Keel, C., Wade, J., Walsh, U., O’Gara, F., and Haas, D. (2000). Small, stable shuttle vectors based on the minimal pVS1 replicon for use in gram-negative, plant-associated bacteria. *Mol. Plant Microbe Interact.* **13**:232–237.
- Itoh, Y., Watson, J.M., Haas, D., and Leisinger, T. (1984). Genetic and molecular characterization of the *Pseudomonas* plasmid pVS1. *Plasmid* **11**:206–220.
- Jack, T., Brockman, L.L., and Meyerowitz, E.M. (1992). The homeotic gene *APETALA3* of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. *Cell* **68**:683–697.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**:816–821.
- Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Tay, T., Newby, G.A., Maiani, J.P., Raguram, A., and Liu, D.R. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat. Biotechnol.* **36**:843–846.
- Labun, K., Montague, T.G., Krause, M., Torres Cleuren, Y.N., Tjeldnes, H., and Valen, E. (2019). CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* **47**:W171–W174.
- Lenk, U., and Sommer, T. (2000). Ubiquitin-mediated proteolysis of a short-lived regulatory protein depends on its cellular localization. *J. Biol. Chem.* **275**:39403–39410.
- Li, J.F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M., and Sheen, J. (2013). Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.* **31**:688–691.
- Liu, H., Ding, Y., Zhou, Y., Jin, W., Xie, K., and Chen, L.L. (2017). CRISPR-P 2.0: an improved CRISPR-Cas9 tool for genome editing in plants. *Mol. Plant* **10**:530–532.
- Maggio, I., Zittersteijn, H.A., Wang, Q., Liu, J., Janssen, J.M., Ojeda, I.T., van der Maarel, S.M., Lankester, A.C., Hoebe, R.C., and Goncalves, M. (2020). Integrating gene delivery and gene-editing technologies by adenoviral vector transfer of optimized CRISPR-Cas9 components. *Gene Ther.* **27**:209–225.
- Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. *Science* **339**:823–826.
- Marillonnet, S., and Grutzner, R. (2020). Synthetic DNA assembly using golden gate cloning and the hierarchical modular cloning pipeline. *Curr. Protoc. Mol. Biol.* **130**:e115.
- Marillonnet, S., Thoeringer, C., Kandzia, R., Klimyuk, V., and Gleba, Y. (2005). Systemic *Agrobacterium tumefaciens*-mediated transfection of viral replicons for efficient transient expression in plants. *Nat. Biotechnol.* **23**:718–723.
- Mikami, M., Toki, S., and Endo, M. (2015). Comparison of CRISPR/Cas9 expression constructs for efficient targeted mutagenesis in rice. *Plant Mol. Biol.* **88**:561–572.
- Mortensen, S., Weaver, J.D., Sathitloetsakun, S., et al. (2019). The regulation of ZCT1, a transcriptional repressor of monoterpene indole alkaloid biosynthetic genes in *Catharanthus roseus*. *Plant Direct* **3**:e00193.
- Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J.D., and Kamoun, S. (2013). Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31**:691–693.
- Nishiguchi, R., Takanami, M., and Oka, A. (1987). Characterization and sequence determination of the replicator region in the hairy-root-inducing plasmid pRiA 4b. *Mol. Genet. Genomics* **206**:1–8.
- Ordon, J., Bressan, M., Kretschmer, C., Dall’Osto, L., Marillonnet, S., Bassi, R., and Stuttmann, J. (2019). Optimized Cas9 expression systems for highly efficient *Arabidopsis* genome editing facilitate isolation of complex alleles in a single generation. *Funct. Integr. genomics* **20**:151–162.
- Ordon, J., Gantner, J., Kemna, J., Schwalgun, L., Reschke, M., Streubel, J., Boch, J., and Stuttmann, J. (2017). Generation of chromosomal deletions in dicotyledonous plants employing a user-friendly genome editing toolkit. *Plant J.* **89**:155–168.
- Rizvi, N.F., Cornejo, M., Stein, K., Weaver, J., Cram, E.J., and Lee-Parsons, C.W.T. (2015). An efficient transformation method for estrogen-inducible transgene expression in *Catharanthus roseus* hairy roots. *Plant cell, Tissue and Organ Culture* **120**:475–487.
- Schultink, A., Qi, T., Lee, A., Steinbrenner, A.D., and Staskawicz, B. (2017). Roq1 mediates recognition of the *Xanthomonas* and *Pseudomonas* effector proteins XopQ and HopQ1. *Plant J.* **92**:787–795.
- Shaul, O. (2017). How introns enhance gene expression. *Int. J. Biochem. Cell Biol.* **91**:145–155.
- Shimada, T.L., Shimada, T., and Hara-Nishimura, I. (2010). A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*. *Plant J.* **61**:519–528.
- Sinapidou, E., Williams, K., Nott, L., Bahkt, S., Tor, M., Crute, I., Bittner-Eddy, P., and Beynon, J. (2004). Two TIR-NB:LRR genes are required to specify resistance to *Peronospora parasitica* isolate Cala2 in *Arabidopsis*. *Plant J.* **38**:898–909.
- Stuttmann, J., Hubberten, H.M., Rietz, S., Kaur, J., Muskett, P., Guerois, R., Bednarek, P., Hoefgen, R., and Parker, J.E. (2011). Perturbation of *Arabidopsis* amino acid metabolism causes incompatibility with the adapted biotrophic pathogen *Hyaloperonospora arabidopsidis*. *Plant Cell* **23**:2788–2803.
- Tsutsui, H., and Higashiyama, T. (2017). pKAMA-ITACHI vectors for highly efficient CRISPR/Cas9-mediated gene knockout in *Arabidopsis thaliana*. *Plant Cell Physiol.* **58**:46–56.
- van der Biezen, E.A., Freddie, C.T., Kahn, K., Parker, J.E., and Jones, J.D. (2002). *Arabidopsis* RPP4 is a member of the RPP5 multigene family of TIR-NB-LRR genes and confers downy mildew resistance through multiple signalling components. *Plant J.* **29**:439–451.
- Wang, Z.P., Xing, H.L., Dong, L., Zhang, H.Y., Han, C.Y., Wang, X.C., and Chen, Q.J. (2015). Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in *Arabidopsis* in a single generation. *Genome Biol.* **16**:144.

- Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S.** (2011). A modular cloning system for standardized assembly of multigene constructs. *PLoS One* **6**:e16765.
- Weigel, D., Alvarez, J., Smyth, D.R., Yanofsky, M.F., and Meyerowitz, E.M.** (1992). LEAFY controls floral meristem identity in Arabidopsis. *Cell* **69**:843–859.
- Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldmann, K.A., and Meyerowitz, E.M.** (1990). The protein encoded by the Arabidopsis homeotic gene *agamous* resembles transcription factors. *Nature* **346**:35–39.
- Ye, X., Williams, E.J., Shen, J., Johnson, S., Lowe, B., Radke, S., Strickland, S., Esser, J.A., Petersen, M.W., and Gilbertson, L.A.** (2011). Enhanced production of single copy backbone-free

transgenic plants in multiple crop species using binary vectors with a pRi replication origin in *Agrobacterium tumefaciens*. *Transgenic Res.* **20**:773–786.

- Zhi, L., TeRonde, S., Meyer, S., Arling, M.L., Register, J.C., 3rd, Zhao, Z.Y., Jones, T.J., and Anand, A.** (2015). Effect of *Agrobacterium* strain and plasmid copy number on transformation frequency, event quality and usable event quality in an elite maize cultivar. *Plant Cell Rep.* **34**:745–754.
- Zhou, H., Liu, B., Weeks, D.P., Spalding, M.H., and Yang, B.** (2014). Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res.* **42**:10903–10914.