the plant journal

SEB Society for Experimental Biology

The Plant Journal (2018) **94,** 735–746 doi: 10.1111/tpj.13893

TECHNICAL ADVANCE

Efficient *in planta* gene targeting in Arabidopsis using egg cell-specific expression of the Cas9 nuclease of *Staphylococcus aureus*

Felix Wolter, Jeannette Klemm and Holger Puchta*

Botanical Institute, Karlsruhe Institute of Technology, POB 6980, 76049, Karlsruhe, Germany

Received 2 January 2018; revised 17 February 2018; accepted 21 February 2018; published online 24 March 2018.
*For correspondence (e-mail holger.puchta@kit.edu).

SUMMARY

Gene targeting (GT), the programmed change of genomic sequences by homologous recombination (HR), is still a major challenge in plants. We previously developed an *in planta* GT strategy by simultaneously releasing from the genome a dsDNA donor molecule and creating a double-stranded break (DSB) at a specific site within the targeted gene. Using Cas9 form *Streptococcus pyogenes* (SpCas9) under the control of a ubiquitin gene promoter, we obtained seeds harbouring GT events, although at a low frequency. In the present research we tested different developmentally controlled promotors and different kinds of DNA lesions for their ability to enhance GT of the acetolactate synthase (ALS) gene of *Arabidopsis*. For this purpose, we used *Staphylococcus aureus* Cas9 (SaCas9) nuclease and the SpCas9 nickase in various combinations. Thus, we analysed the effect of single-stranded break (SSB) activation of a targeted gene and/or the HR donor region. Moreover, we tested whether DSBs with 5′ or 3′ overhangs can improve *in planta* GT. Interestingly, the use of the SaCas9 nuclease controlled by an egg cell-specific promoter was the most efficient: depending on the line, in the very best case 6% of all seeds carried GT events. In a third of all lines, the targeting occurred around the 1% range of the tested seeds. Molecular analysis revealed that in about half of the cases perfect HR of both DSB ends occurred. Thus, using the improved technology, it should now be feasible to introduce any directed change into the Arabidopsis genome at will.

Keywords: *Arabidopsis thaliana*, double-stranded break repair, engineered nucleases, genome editing, homologous recombination, targeted mutagenesis.

INTRODUCTION

It has been known for more than 20 years that gene targeting (GT) by homologous recombination (HR) (Puchta *et al.*, 1996), or gene knockout by non-homologous end-joining (NHEJ) (Salomon and Puchta, 1998), can be achieved by the induction of double-stranded breaks (DSBs) with sequence-specific endonucleases in plant genomes. However, it was only with the development of the CRISPR/Cas technology that tools became available that can be used with ease in any molecular biology laboratory, with new applications of the technology currently transforming plant biology (Schiml and Puchta, 2016; Mahas *et al.*, 2017; Mahfouz, 2017; Malzahn *et al.*, 2017; Puchta, 2017). Using Cas9 nucleases for trait improvement by knocking out specific genes by NHEJ, the technology has become routine for applications in agriculture (Baltes *et al.*, 2017; Pacher and Puchta, 2017; Scheben *et al.*, 2017; Yin *et al.*, 2017).

Unfortunately, despite DSB induction, the frequencies for controlled genomic change by homologous recombination (HR) are still disappointing. Multiple studies performed with different synthetic nucleases (Wright *et al.*, 2005; D'Halluin *et al.*, 2008; Shukla *et al.*, 2009; Townsend *et al.*, 2009; Ayar *et al.*, 2013; de Pater *et al.*, 2013; Watanabe *et al.*, 2016) showed that the GT frequencies were low, often below the 1% range. The reason for this is that, in somatic plant cells, NHEJ is much more efficient than HR in repairing DSBs (Knoll *et al.*, 2014). Using Cas9-mediated induction of DSBs, several strategies have been developed to enhance HR-mediated GT frequencies in plants (for a recent review see Steinert *et al.*, 2016). The group of Dan

Voytas, for instance, developed an innovative geminivirus replicon-based strategy which enabled them to improve GT frequencies in some plant species (Baltes *et al.*, 2014; Čermák *et al.*, 2015). However, at least for monocots, no fertile plants have been reported to be obtained with this technology, as yet (Gil-Humanes *et al.*, 2017; Wang *et al.*, 2017). The use of DNA oligonucleotides (Svitashev *et al.*, 2015; Sauer *et al.*, 2016) has also been applied successfully for gene correction in plants. In one attempt, a chimeric sgRNA was used as a template for DSB repair (Butt *et al.*, 2017). Additionally, inhibition of NHEJ (Endo *et al.*, 2016; Nishizawa-Yokoi *et al.*, 2016) can help to improve GT frequencies. Nevertheless, although progress has been achieved, in most cases the frequencies of GT are low and further improvements are needed.

We chose a different strategy to overcome the limitations of GT experiments in plants. Due to the low efficiency of GT after DSB induction, a reasonable transformation frequency is prerequisite for GT experiments. Unfortunately, many crop plants are hard to transform and large amounts of plant material have to be processed *in vitro*. The regeneration of edited cells into fertile plants presents an additional challenge (Altpeter *et al.*, 2016). Having these limitations in mind, we developed the so-called *in planta* GT technology (Fauser *et al.*, 2012), which relies on a chromosomal donor molecule, e.g. a stably integrated T-DNA, carrying the HR donor sequence as well as the Cas9 expression cassette. Cas9 expression leads to the simultaneous induction of three DSBs. The HR donor molecule is excised out of the genome at the same time as a DSB is induced at the target locus to enhance HR (for principle see Figure 1). By expressing the nuclease under a constitutive promoter, GT can occur during plant development. The important factor is the transfer of the targeting events in the germline to the next generation. By simply screening seeds or seedlings, clonal GT events can be identified. Thus, large amounts of *in vitro* culture for obtaining hundreds of transformation events could be avoided. We adopted the technology for the use of the Cas9 nuclease of *S. pyogenes* (SpCas9) under the control of a ubiquitin gene promoter in *Arabidopsis* and were able to obtain GT events in the range of 1 out of 700 seeds (Schiml *et al.*, 2014).

As the frequencies were low, we were eager to find ways to improve the system. We were able to recently show by comparing the efficiency of different Cas9 orthologues for mutation induction in Arabidopsis that the protein from *S. aureus* (SaCas9) was more efficient than SpCas9 in obtaining NHEJ-induced mutation frequencies (up to over 90%, Steinert *et al.*, 2015). Moreover, recent studies indicated that in Arabidopsis the expression control by developmentally regulated promoters such as an egg-specific promoter (Wang *et al.*, 2015) or a reproductive tissue-associated promoter (Yan *et al.*, 2015) might also enhance mutation frequencies.

By using a mutated form of SpCas9 in which the open reading frame was changed by a single amino acid in the active site of one domain – making a nickase out of the nuclease – we could show that intrachromosomal HR can be induced to an even higher extent by a SSB than by a DSB (Fauser *et al.*, 2014). Furthermore, by the use of two sgRNAs, leading the SpCas9 nickase to cutting sites within close proximity of each other on opposite strands, 3′ and 5′ overhangs can be obtained (Schiml *et al.*, 2016). As the production of 3′ intermediates is an important step during various HR reactions (Knoll *et al.*, 2014), we speculated that the direct induction of 3′ overhangs might be a more efficient mediator of GT. We have shown previously that SpCas9 and SaCa9 do not interfere with each other's sgRNAs in plant cells (Steinert *et al.*, 2015). Thus, application of both enzymes in the same cell can be used to induce SSBs and DSBs at different sites, simultaneously (Puchta, 2016).

Having this set of new tools in hand, we tested if any combination of them would be able to sufficiently enhance *in planta* GT efficiencies to make the technology feasible for daily use for the plant community.

RESULTS

Comparison of intrachromosomal HR induction rates by SpCas9 and SaCas9 nucleases and nickases

As we wanted to test the influence of the induction of SSBs and the simultaneous induction of SSBs and DSBs on *in planta* GT, we needed to define which Cas9 nuclease and nickase orthologues would be most appropriate for this application. First, we compared the Cas9 nucleases and nickases of *S. aureus* and *S. pyogenes* for their potential to induce intrachromosomal HR in the Arabidopsis reporter line IU.GUS (Orel and Puchta, 2003; Puchta and Hohn, 2012). This reporter line contains a stably integrated, homozygous, β-glucuronidase gene (*uidA*, GUS) disrupted by a short spacer (Figure 2(a)). Upon DSB induction in the spacer, functional restoration of the GUS gene can be achieved via the synthesis-dependent strand annealing (SDSA) pathway using an inverted homology located upstream of the GUS gene. We used plant codon-optimized clones of the SpCas9 nuclease and nickase (Fauser *et al.*, 2014), as well as the SaCas9 nuclease (Steinert *et al.*, 2015), as described before. We obtained a clone of the SaCas9 nickase by mutation of aspartic acid to alanine at position 10 within the RuvC domain of the SaCas9 ORF (Steinert *et al.*, 2015). This mutation is analogous to the one used for production of the SpCas9 nickase (Fauser *et al.*, 2014). We combined the four enzymes with sgRNAs targeting the spacer region of the reporter line. After floral dip transformation of the constructs into the reporter lines, we picked primary transformants and quantified GUS activity via a 4MUG-assay. As described previously (Fauser

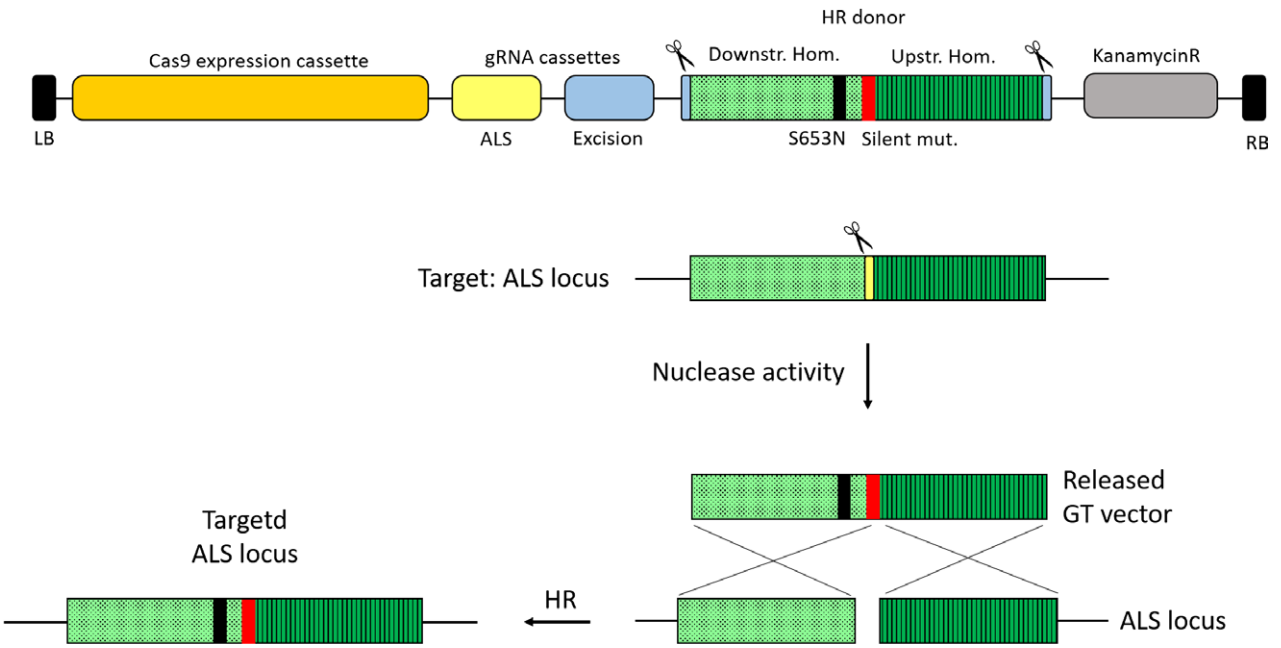

Figure 1. Outline of the *in planta* GT approach as applied for the induction of a point mutation in the ALS gene of *Arabidopsis*.
The T-DNA construct that is transformed into *Arabidopsis* plants consists of the SaCas9 nuclease driven by the respective promoter (in orange), two sgRNA expression cassettes, a HR donor (in green) and a transformation marker (in grey). One sgRNA targets the ALS (in yellow), the other recognizes a sequence flanking the HR donor (in blue), which harbours a total of 1542 bp homology containing the S653N point mutation and silent mutations at the cutting site. DSB induction leads to the activation of the target site and simultaneous release of the HR donor molecule, which can then can be used as template for repair of the target locus by HR.

et al., 2014), the nickase induced HR more efficiently than the nuclease in the case of SpCas9 (Figure 2(b)). Surprisingly, such a difference was not found for SaCas9. Moreover, the induction of HR by the nickase of *S. pyogenes* was about twice as efficient than with the nickase of *S. aureus*. Thus, we chose the combination of the most efficient nuclease (*S. aureus*) and the most efficient nickase (*S. pyogenes*) for employment in the GT experiments that followed.

Set up of the *in planta* gene targeting assay

To test whether new approaches might improve GT frequencies in Arabidopsis, we chose the acetolactate synthase gene (*ALS*, AT3G48560) as target. This enzyme is essential as it is required for the synthesis of branched chain amino acids. It is inhibited by sulfonylurea and imidazolinon herbicides. However, several point mutations rendering the enzyme resistant against specific herbicides are known (Lee *et al.*, 1988). One of these, used before by Endo *et al.* (2006), is the exchange of a serine for an asparagine at position 653 (S653N), conferring resistance to imazapyr (IM). Therefore, we included the respective point mutation in the HR donor sequence. This sequence included 800 bp of homology on one side and 742 on the other side of the break, which was located 120 bp from the S653N point mutations (Figure 1). The homology arms are flanked by *S. aureus* target sequences including a NNGGGT PAM. This set up could be used simultaneously for SSB induction by SpCas9-D10A, as its PAM (NGG) is part of the same sequence motive. To avoid cutting the HR donor, we introduced four more silent point mutations spanning the sgRNA and PAM sequence within the repair template (see Figure S3).

In the first set of experiments, the Cas9 open reading frames were expressed by the constitutive PcUbiquitin4-2 promoter from *Petroselinum crispum* (PcUbi4-2). In all experiments, the sgRNAs were set under the control of the Arabidopsis U6-26 promoter. We designed seven constructs to copy the S653N point mutation into the genomic ALS open reading frame (Figure 1). One approach is analogous to our previously reported approach for ADH1 GT (Schiml *et al.*, 2014), inducing a DSB for induction of HR in the target (ALS) as well as for excision of the HR donor, but employs the SaCas9 nuclease instead of the SpCas9 nuclease. Four approaches aimed to induce SSBs instead of DSBs either in the template or the HR donor or both. In addition, we used the SpCas9 nickase for generating 50 bp 5′ and 3′ overhangs spanning the S653N site and adapted the silent mutations within our HR donor sequence accordingly (for details see Figure S3). In approaches using both SaCas9 and SpCas9-D10A simultaneously, we co-transformed two T-DNAs with different selectable markers to keep T-DNAs within a manageable size. In case we released a single-stranded HR donor, we tested two

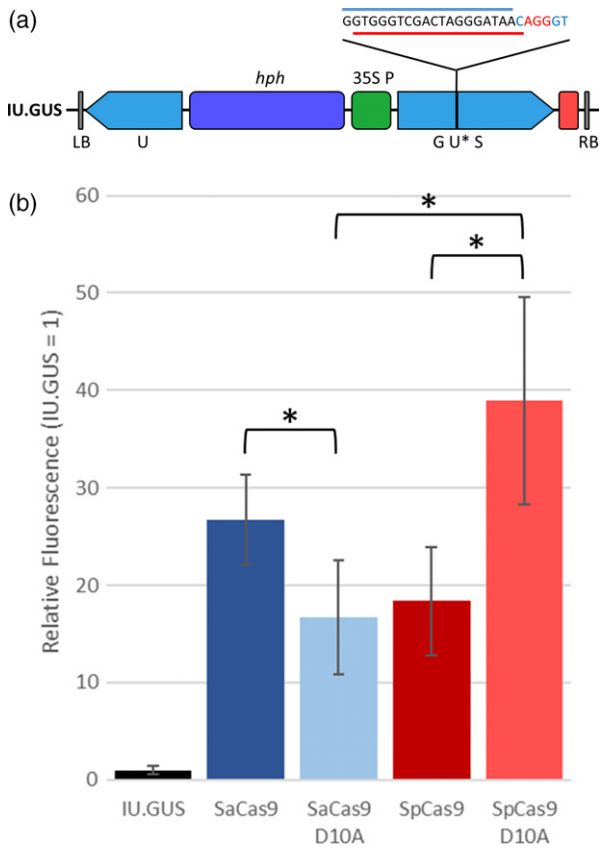

Figure 2. Efficiency of HR induction by the SaCas9 and SpCas9 nucleases and nickases as quantified by a fluorescent assay using a transgenic SDSA reporter line.
(a) The SDSA-based reporter line (IU.GUS) consists of overlapping parts of the β-glucuronidase gene. Induction of a lesion at the indicated target sites leads to the restoration of the GUS ORF by HR. The SaCas9 and SpCas9 target sequences are depicted above the chart, indicated by a blue and red bar, respectively. The SpCas9 PAM is in red and completely contained within the SaCas9 PAM ('CAGGGT'), whose additional nucleotides are in blue.
(b) Lesion- and enzyme-dependent HR induction as measured by activity of restored GUS ORF via 4MUG-assay. Values are given in relation to a control (IU.GUS) without lesion induction. Asterisks (*) indicate *P*-values of <0.05 for two-sample *t*-tests, one-sided with equal variance.

different target sites for nick induction in the target, one on the strand complementary to the donor and one on the donor-like strand.

The constructs were transformed into Arabidopsis via the floral dip technique and at least 40 primary transformants per construct were selected for further cultivation using their resistance to the transformation marker(s). We observed impairments in plant vitality in the T1 generation for the approaches involving DSB induction in ALS, but not for approaches involving SSB induction, reflecting NHEJ mutagenesis by DSBs which is absent for SSBs. To test whether efficient DSB induction occurred at the target, we amplified a polymerase chain reaction (PCR) fragment containing the target site from 10 T1 plants transformed with the SaCas9 nuclease and performed Sanger sequencing. As shown in Figure 3 as an example, the sequencing chromatograms showed a pattern of multiple overlapping bands, beginning exactly at the DSB site. This indicates efficient DSB induction by the nuclease followed by NHEJ repair.

To screen for GT events by HR, the seeds of the T1 lines were harvested separately and screened on imazapyr (IM)-containing medium. In the course of our analysis, we determined how many of the independently transformed lines produced resistant seedlings whilst also determining how many resistant seedlings could be isolated per positive line (Table 1).

Surprisingly, we were not able to detect any targeting events when an SSB was introduced in the genomic target locus independently of whether a nuclease or nickase was used to release the HR donor from the chromosomally-integrated T-DNA construct. Additionally, the approaches producing 3′ or 5′ overhangs resulted in hardly any or no resistant plants (Table 1). Only DSB induction of the genomic target locus lead to a significant number of resistant plants. If the approach was combined with SSBs for HR donor release, 5 out of 28 lines (18%) produced resistant plants and in terms of excising the HR donor via DSBs the ratio was higher (27%; 16 of 60 lines). In one of these lines (DSB/DSB, UBI #14), the targeting event occurred at or early after transformation, as practically all of the tested progeny were resistant (1580 resistant plantlets, not included in the 'resistant plants (T2)' column but referred to as 'GT T1 lines' in Table 1). When not taking this event into account, 56 resistant seedlings were obtained in total. Thus, in respective lines, about 1 out of 300 seeds was resistant, which is higher than the frequencies we obtained with SpCas9 before (1 out of 700; Schiml *et al.*, 2014), indicating that SaCas9 is not only more efficient for NHEJ mediated mutation (Steinert *et al.*, 2015), but also for HR-mediated GT in Arabidopsis.

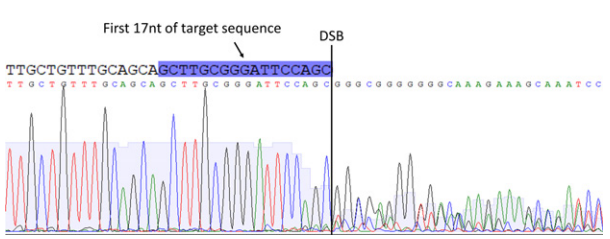

Figure 3. SaCas9 nuclease activity as measured by the induction of NHEJ at the target site.
DNA was extracted from T1 plants transformed with SaCas9 and the sgRNA targeting the *ALS* gene. The target locus was amplified and subjected to Sanger sequencing. The chromatogram shows multiple overlapping sequences starting at the site of DSB induction, indicating NHEJ-mediated mutagenesis at high level.

Table 1 Quantitative outcome of the different GT approaches

Induction in target	GTV excision	Promoter	Amount of T1 lines	Amount of seeds (T2, estimate)	GT T1 lines	GT positive lines	Positive line rate (%)	IM-resistant plants (T2)	Mean GT rate (T2) (%)
DSB	DSB	PcUbi4–2	60	~70 k	1	16	27	56	0.3
DSB	SSB	PcUbi4–2	28	~40 k	0	5	18	8	0.14
SSB (+)	SSB	PcUbi4–2	39	~70 k	0	0	0	0	0
SSB (+)	DSB	PcUbi4–2	40	~60 k	0	0	0	0	0
SSB (−)	SSB	PcUbi4–2	40	~80 k	0	0	0	0	0
3′overhangs	DSB	PcUbi4–2	33	~20 k	0	1	3	1	0.2
5′overhangs	DSB	PcUbi4–2	35	~30 k	0	0	0	0	0
DSB	DSB	AtCLV3	16	~30 k	0	0	0	0	0
DSB	DSB	AtYAO	8	~9 k	0	3	38	3	0.08
DSB	DSB	AtEC1.1/1.2	74	~140 k	1	55	74	952	0.97
DSB	SSB	AtEC1.1/1.2	40	~70 k	0	18	45	56	0.19
SSB (+)	DSB	AtEC1.1/1.2	40	~100 k	0	2	5	3	0.05
SSB (+)	SSB	AtEC1.1/1.2	40	~80 k	0	3	8	4	0.07
SSB (−)	SSB	AtEC1.1/1.2	40	~90 k	0	5	13	21	0.22
3′overhangs	DSB	AtEC1.1/1.2	40	~70 k	0	3	8	4	0.06
5′overhangs	DSB	AtEC1.1/1.2	40	~60 k	1	21	53	219	0.65

The different experimental approaches are given by the first three columns of the table. The first column specifies the kind of DNA lesion induced in the target, the second column specifies if a DSB or SSB is used for release of the HR donor. SSB (+) or (−) indicates whether the SSB was induced in the sense or antisense strand of the ALS ORF in the target locus. The column 'GT positive lines' gives the number of lines showing heritable GT events in the T2 generation. The column 'GT T1 lines' indicates the number of lines with T0 or early T1 GT events giving rise to solely IM-resistant T2 plants. The mean GT rate was calculated as follows: first the GT rate for each single GT positive line was determined by dividing the number of resistant T2 plants by the total number of T2 seeds. Then the mean GT rate was calculated by averaging the obtained GT rates of all GT positive lines. The abundant resistant plants from 'GT T1 lines' are excluded from the calculation, as the high numbers of resistant plants (all progeny) would distort the statistics. As the GT numbers are based on IM-resistant seedlings, they contain perfect as well as ectopic GT events.

Efficient *in planta* GT using egg cell-specific expression of Cas9

As we did not find any indication that the induction of SSBs is superior compared with that of DSBs for *in planta* GT, we next tested whether tissue-specific promoters for Cas9 expression can enhance GT efficiency in our system. We selected the promoters of the stem cell identity regulating CLAVATA3 (CLV3, AT2G27250) (van Ex *et al.*, 2009), the YAO gene (AT4G05410) which showed strong expression in meristematic tissue (Yan *et al.*, 2015), and the egg cell-specific EC1.1 (AT1G76750, Sprunck *et al.*, 2012). The EC1.1 promoter was fused with the EC1.2 enhancer and combined with the terminator of *Pisum sativum* rbcS E9 [encoding the small subunit (rbcS) of ribulose-1,5-bisphosphate carboxylase] which was found to strongly enhance efficiency of egg cell-specific Cas9 expression (Wang *et al.*, 2015). The promoters were combined with the DSB/DSB approach and analysed for GT efficiency as described above. For CLV3 and YAO promoters, many primary transformants died or were sterile, indicating strong meristematic ALS mutagenesis. The surviving lines from these promoters yielded no improvement in GT frequency (Table 1). Interestingly, the egg cell promoter (EC1) performed extremely well: 55 out of 74 lines (74%) generated heritable GT events. Again, as for the UBI promoter we found one line with completely resistant progeny (899 plantlets, line DSB/DSB, EC #b5) indicating an early event at, or shortly after, transformation. But even with the exclusion of this line, we obtained 952 resistant plants, corresponding to a mean GT frequency of 1% over all positive lines. The superiority of egg cell-specific Cas9 expression became more pronounced when GT frequencies of single lines were considered (Figure 4). For the EC1 promoter, around two-thirds of the lines were above the 0.5% GT rate, more than one-third were above 1%, and individual lines showed up to 6% GT frequency (excluding the line showing 100% GT frequency). For the ubiquitin gene promoter, only two lines reached above 1% efficiency apart from the line showing 100% GT (Figure 4).

Having found an efficient promoter for our GT system, we next tested again whether under these conditions, induction of staggered DSB ends might enhance GT efficiency. In addition, we readdressed the approaches involving the induction of SSBs in the targeted chromosomal gene and for HR donor excision in combination with the EC1 promoter. In this experiment, we were able to generate GT events for all approaches tested (Table 1). However, frequencies of the approaches involving single nicks in the target remained below the DSB/DSB approach. This includes the DSB/SSB approach involving SSB for excision

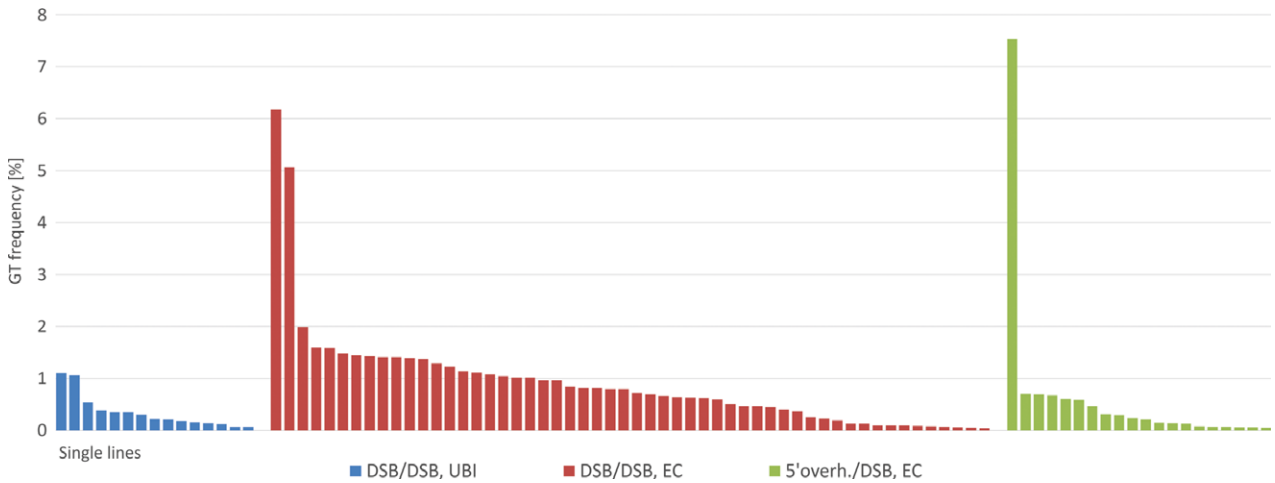

Figure 4. GT frequencies of single lines of selected approaches, DSB/DSB, UBI in blue; DSB/DSB, EC in red and 5′ overhang/DSB, EC in green (meaning 5′ overhangs in the target and DSBs for HR donor release).
The GT frequencies of individual GT positive lines are depicted. The numbers include perfect as well as ectopic GT events (for details on GT rate calculation see Table 1). The DSB/DSB, EC approach emerges as the most efficient and most consistent approach, with a majority of the lines reaching the 1% range.

of single-stranded repair template(s), indicating that DSBs for HR donor excision are superior, as was already indicated from the experiments with constitutive Cas9 expression. Regarding staggered breaks, we found induction of 5′ overhangs was a much more potent mediator of GT events than 3′ overhangs. For the 5′ overhang approach, 53% of the lines yielded GT events, and again one of these apparently occurred in egg cells or early zygotes soon after transformation (5′/DSB, EC #19) producing only resistant plantlets (1704 in number). The mean GT of the other lines was around 0.6%.

Molecular analysis of recombination events

The production of a functional marker alone does not prove that the GT event was due to perfect HR. Indeed, often besides perfect GT events, the break can be repaired by a combination of HR and NHEJ at the target locus or due to copying processes of the homology from the locus to the HR donor with concomitant integration of the donor elsewhere in the genome, leaving the targeting site undisturbed ('ectopic targeting'; Puchta and Fauser, 2013). To discriminate between these possibilities, we performed molecular analysis of a large number of resistant seedlings. Because we only introduced subtle changes in the target locus, we could not perform Southern blot analysis to unambiguously prove perfect GT. Therefore, we performed PCR followed by Sanger sequencing over the target locus.

We amplified the ALS locus using primers outside the homology region to prevent amplification of HR donor sequences from the integrated T-DNA (Figure 5). In total, we analysed 171 IM-resistant plants across all approaches. Of these, 88 showed perfect GT (51%) including the intended S653N and silent mutations as well as correct junctions (see Figure 6 for an example of a perfect targeting event). Here, 20 of the 86 plants tested positive were already biallelic for the GT allele. Monoallelic and biallelic plants were differentiated by the chromatogram showing a mixed signal or only the GT allele signal, respectively. We found for all approaches that resulted in resistant plants – independent of the promoter or the DNA lesion applied – perfect targeting events. The molecular analysis is summarized in Table 2 (for details on individual lines, see Table S1). In addition, we tested several perfect GT events for segregation. They all behaved as expected: progeny of biallelic GT plants remaining biallelic, whereas heterozygous GT plants showed 3:1 segregation of the GT allele. We checked several of the GT positive plants for presence of the transgene and identified two transgene-free homozygous GT plants (DSB/DSB, UBI #13-1 and DSB/DSB, UBI #13-4).

However, it is noteworthy that different targeting event classes were not equally distributed between events. Whereas for the SaCas9 nuclease-based targeting, at least in half of the cases (36 of 72 in the case of the egg cell and 29 of 39 in the case of the ubiquitin gene promoter) perfect GT could be achieved, in case of the 5′ overhangs only 4 out of 34 events turned out to have a perfectly modified *ALS* gene.

We hypothesized that ectopic targeting, the copying of sequence information from the native ALS to the HR donor, is responsible for IM resistance in plants not showing perfect GT. If the missing N-terminal part of the ALS and parts of the promoter are copied to the HR donor, a completely functional, IM resistance-mediating ALS cassette would be generated (Figure S1). To confirm this

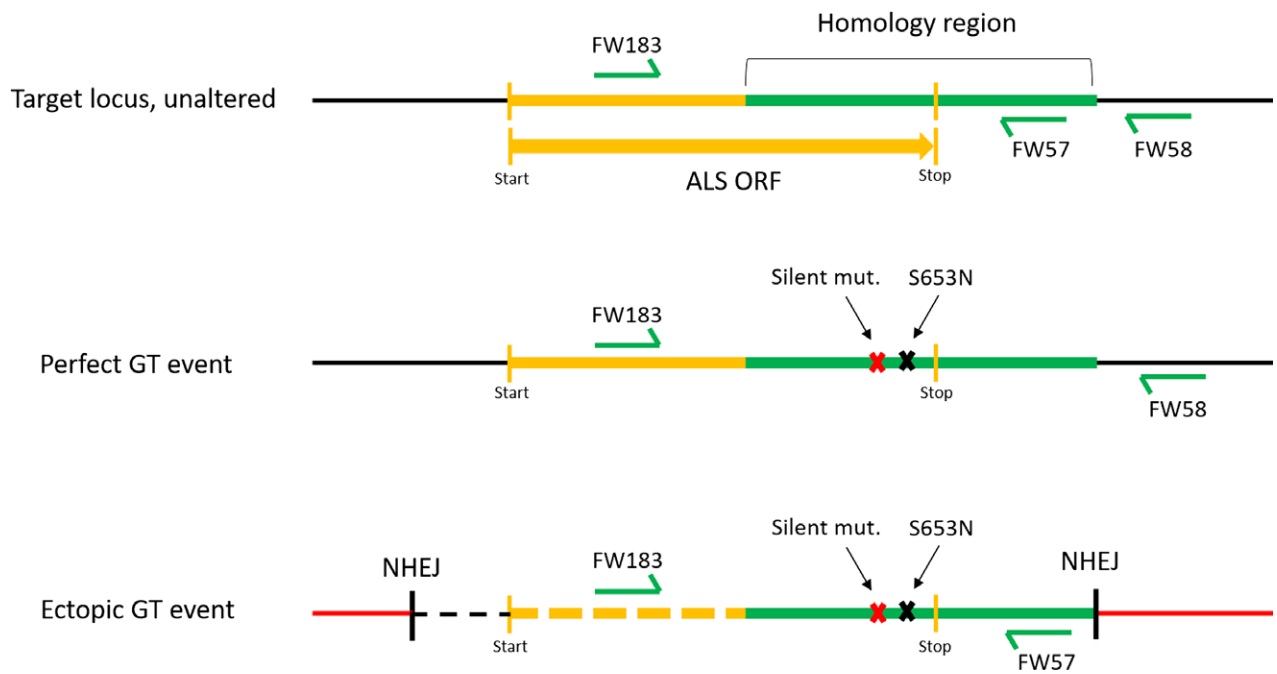

Figure 5. Molecular characterization of GT events by sequencing over the target locus.
The primer combination FW183/FW58 binding completely outside the homology regions of the ALS locus was used to verify perfect gene targeting events by the presence of the S653N point mutation, the silent mutations and both correct junctions on one single amplicon. The primer FW57 binds inside the homology region and, together with FW183, will amplify not only native and perfectly targeted ALS genes but also ectopic GT events. Presence of the S653N and silent mutations on this amplicon together with a lack of amplification of the same mutations with FW183/FW58 primer pair were classified as ectopic GT event. For an explanation on ectopic GT events, see Figure S1.

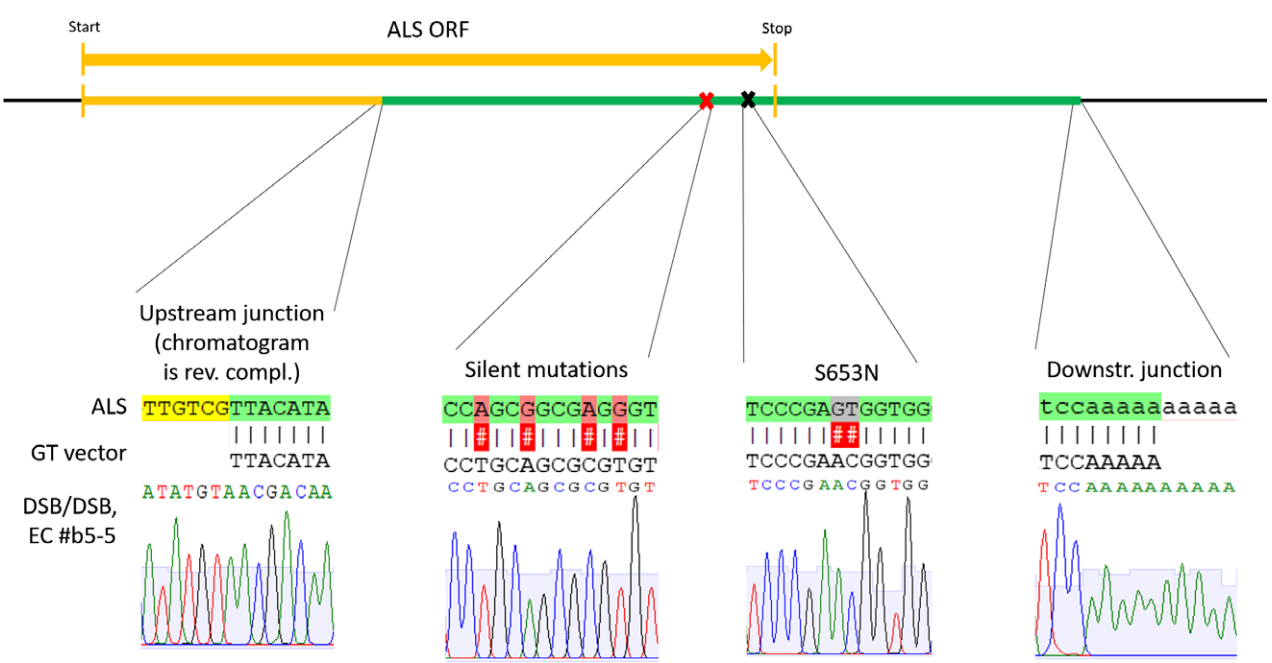

Figure 6. Example of the molecular structure of a perfect GT event.
The chromatograms of a homozygous GT plant (DSB/DSB EC 5-5) are aligned with the native sequence and its modified part on the HR donor (labelled 'GT vector'). The chromatogram shows perfect match with the HR donor, not the native sequence. For the upstream junction on the left, in contrast to the three other chromatograms, the sequencing reaction was performed with a reverse primer, so the chromatogram is the reverse complement of the aligned sequences. Heterozygous GT plants were identified by showing overlaid signals from both the GT and WT allele at the modification sites.

Table 2 Molecular analysis of representative GT events

	Total analyzed	Perfect GT	Rate in %
DSB/DSB, UBI	39	29	74
DSB/SSB, UBI	8	7	88
3′ overh./DSB, UBI	1	1	100
DSB/DSB YAO	2	1	50
DSB/DSB EC	72	36	50
DSB/SSB EC	2	2	100
SSB+/DSB EC	1	1	100
SSB+/SSB EC	4	4	100
SSB−/SSB EC	6	1	17
5′ overh./DSB EC	34	4	12
3′ overh./DSB EC	2	2	100
Total	171	88	51

hypothesis, we used another primer combination specifically designed to amplify such events, where one primer binds to the expected upstream elongation of the HR donor whereas the other primer binds in the distal homology region to provide amplification independent from integration site. As one primer binds outside the homology region, the native HR donor cannot be amplified. The native ALS locus will always be co-amplified, but wild-type (WT) plants can be distinguished from ectopic GT plants by not additionally showing S653N and silent mutations in Sanger sequencing of the amplified junction. We analysed 40 non-perfect GT events and classified 39, due to the respective expected pattern, as ectopic GT events. In addition, we found in these IM-resistant T2 seedlings in many cases homozygous mutations destroying the ALS ORF (Figure S2). If not for ectopic GT events, such plants would lack a functional ALS and would not be able to survive.

To sum up, our molecular analysis demonstrated that the *in planta* GT strategy using the SaCas9 nuclease under egg cell specific promoter can result in perfect GT events within the range of 1%.

DISCUSSION

We recently adopted the *in planta* GT strategy (Fauser *et al.*, 2012) for use with the CRISPR/Cas system (Schiml *et al.*, 2014). We applied SpCas9 under the control of a ubiquitin gene promoter and were able to produce seedlings with GT events at a frequency of about one targeting event out of 700 seedlings, independent of whether the event was due to perfect or one-sided HR. This was a disappointing low frequency, too small to be useful for most plant molecular biology laboratories. In the current communication, we addressed three different parameters by which alone or in combination the efficiency of the procedure might be improved: The nature of the lesion inducing enzyme, the nature of the DNA lesion itself and the pattern of expression of the enzyme during plant development.

Besides the 'classical' Cas9 nuclease from *S. pyogenes* (Fauser *et al.*, 2014), we also adapted other bacterial Cas9 nucleases for use in plant genome engineering. In the course of these studies, we discovered that at least in the case of Arabidopsis the use of Cas9 from *S. aureus* in combination with specific PAM sequences leads to much higher mutation frequencies based on DSB-induced NHEJ (10 to 30% in comparison with 60 to almost 100%). Therefore, we expected that the replacement of the SpCas9 by the SaCas9 nuclease would improve *in planta* GT frequencies. Comparing the data of our previous work (Schiml *et al.*, 2014) with the data obtained here (in both cases using the same ubiquitin gene promoter) indicates that this is indeed the case (1 out of about 300 instead of 1 out of 700 seedlings). Thus, at least for Arabidopsis, it is generally advisable, and if due to its more restricting PAM requirements feasible, to use the SaCas9 nuclease instead of the SpCas9 nuclease for DSB induction, irrespective of whether HR- or NHEJ-based changes are desired.

It has been demonstrated that Cas9 nucleases can be converted into nickases simply by inducing a mutation in one or the other active centres of the enzyme, which both independently induce an SSB (Jinek *et al.*, 2012). Interestingly, in respect to efficiency, nickases and nucleases differed in our experiments in their efficiencies of HR induction, if we compare the proteins of both organisms. As shown in Figure 2, we found that the nickase of SpCas9 outcompetes that of SaCas9 by a factor of two, in the induction of intrachromosomal HR. Although we have no easy explanation for this result, it was fortunate for our experiments that we could combine the best available nuclease (SaCas9) with the best available nickase (SpCas9). Both enzymes could be used simultaneously within the same approach (Puchta, 2016), as we were able to show that their respective sgRNAs do not interfere with each other in Arabidopsis cells (Steinert *et al.*, 2015).

For the purpose of gene editing by NHEJ, the nickase was used in conjunction with two sgRNAs that recognized sequences on opposite strands 30 to 50 bp apart to induce paired nicks, resulting in a DSB with a longer 3′ or 5′ overhangs. Therefore, the specificity of DSB induction could be dramatically increased by enlarging the recognition site from 23 to 46 nt (Mali *et al.*, 2013). With the setup of a SpCas9 nickase for plants (Fauser *et al.*, 2014), we gained a tool that enabled us to induce not only SSBs instead of DSBs, but also DSBs with single-stranded (ss) overhangs (Schiml *et al.*, 2014) – in contrast with DSBs with blunt ends as produced by the SpCas9 nuclease. We were hopeful that both kinds of approaches might improve *in planta* GT efficiency. Using transgenic HR reporter constructs (Puchta and Hohn, 2012) in previous experiments, we found indications that SSBs are more efficient than DSBs in inducing intrachromosomal recombination (Fauser *et al.*, 2014). Therefore, we tested whether the induction of

SSBs in the *ALS* gene would enhance GT efficiency. In addition, we speculated that two flanking nicks on the same strand might lead to the release of a single-stranded HR donor, leaving behind a single-stranded gap that could again be filled in by DNA synthesis. This scenario would enable multiple excision events, resulting in multiple copies of single-stranded repair templates for HR which might enhance efficiency. Accordingly, we also tested whether the induction of two SSBs for HR donor excision improves GT efficiency. The two approaches were also combined by simultaneously using SSB induction for both HR donor and template. As shown in Table 1, when using two different promoters, instead of an enhancement using SSBs we obtained a dramatically smaller number of resistant seedlings in most cases. It was only for DSB induction in the ALS gene whereby frequencies were moderately higher. Nevertheless, with both promoters used in this study, DSBs immediately adjacent to both sides of the HR donor for its release outperformed SSBs in respect to the percentage of positive lines as well as in respect to the number of resistant seedlings, by two to 10 times. This demonstrates that although it is helpful to cut the HR donor out of the genome using a nuclease, the most important step for initiating HR is the induction of a DSB within the target locus. Indeed, we showed some time ago that ectopically integrated sequences can be used for GT without cutting them out of the genome, although at a low efficiency (Puchta, 1999). In line with our results, the group of Dan Voytas recently reported that they found no indication that induction of SSBs in the target locus would outcompete DSBs in replication associated GT (Čermák *et al.*, 2017).

As SSBs, in contrast to DSBs, are not repaired by the error prone NHEJ which excludes all mutated junctions from re-cutting, we expected that SSBs would be especially well suited for GT approaches. The explanation as to why this is not the case may be due to the differing mechanisms of DSB-induced and SSB-induced HR. Over the years, considerable evidence was accumulated in multicellular eukaryotes, including plants, that indicated that DSB-induced GT is best described by the SDSA mechanism of recombination (Puchta, 1998; Paix *et al.*, 2017). We were able to define key factors involved in SDSA in plants (Roth *et al.*, 2012). Therefore, it will be interesting to test with the same approach whether the same proteins are also required for SSB repair via HR. Conflicting results were reported from mammalian cells whether SSB- and DSB-induced HR share the same key factors (Davis and Maizels, 2014; Vriend *et al.*, 2016).

An essential conserved step in HR in all eukaryotes is the resection of the double-stranded DNA ends to longer ss 3′ overhangs (or review see Daley *et al.*, 2015). It was therefore temping to speculate that the production of a DNA structure mimicking such an intermediate at the target locus, by the use of the paired nickase approach, might boost GT efficiencies. We had previously set up the paired nickase approach for plants (Schiml *et al.*, 2014) and were able to demonstrate that the induction of this kind of lesion can induce NHEJ efficiently, leading to not only deletions but also to tandem duplication within the target site (Schiml *et al.*, 2016). Unfortunately, application of this approach did not result in higher GT frequencies than with the use of the nuclease, independent of the promoter used (see Table 1). Even more surprising was the fact that the production of 5′ ends resulted in more IM-resistant plants than with the 3′ends, at least with the egg cell specific promoter. Interestingly, a superiority of 5′ over 3′ overhangs on GT efficiencies has been reported for mammalian cells before (Mali *et al.*, 2013; Bothmer *et al.*, 2017). A similar result was also recently found in transient assays in tobacco cells (Čermák *et al.*, 2017), although no GT events were regenerated and characterized on the molecular level. Our study demonstrates the importance of molecular characterization for the evaluation of GT events, as for most of the seedlings the IM resistance was due to ectopic and not due precise GT in the case of the 5′ overhang approach (30 out of 34 tested). Consequently, the induction of 5′ overhangs cannot be regarded as a feasible alternative to the induction of a DSB for *in planta* GT.

The use of specific promoters for Cas9 expression turned out to be crucial for the production of heritable mutations in Arabidopsis. In the early days, little success was achieved using CaMV 35S promoters. We were able to improve this efficiency by the use of a ubiquitin gene promoter (Fauser *et al.*, 2014), and subsequently with the use of an egg cell specific gene promoter whose expression was earlier reported to be efficient for obtaining NHEJ-based mutants (Wang *et al.*, 2015). The decisive factor seems to be to enable Cas9 expression at a high level in cells that contribute to the germline. Based on these observations, it became important to test a number of developmentally regulated promoters to determine if they could improve GT efficiency. Results showed that an egg cell-specific promoter worked best for improving *in planta* GT – more than by any other means – by enhancing the number of positive lines and also the number of resistant plants per line by more than a factor of three, on average.

Our molecular analysis revealed that in almost all tested cases, the restoration of IM resistance was due to a targeting event based on HR. In principle, not every targeting event has to be due to the repair of both ends of the DSB by HR using the donor as template. Two other possible outcomes were envisaged (Puchta and Fauser, 2013). The first being one-sided events, by which only one end of the DSB is repaired by HR and the other end is repaired by NHEJ. This is a phenomenon that was recognized even in the earliest studies of DSB-induced GT in plants (Puchta *et al.*, 1996; Puchta, 1998). Conversely, the copying process

cannot only occur into the genome at the site of the DSB, but also from this genomic site onto the HR donor that then integrates elsewhere in the genome ('ectopic targeting', see Figure S1). Due to the selection procedure applied in this study, we only found ectopic GT events besides perfect events. One-sided events escaped selection because by NHEJ the ALS ORF would lose its function, instead of gaining resistance. For most setups (Table 2 and Table S1), we found perfect and ectopic targeting at similar frequencies, it was only for the 5′ overhangs that ectopic targeting outnumbered perfect events by a factor of more than 8. This indicates that a free genomic 5′ end might especially favour the copying process onto the free 3′ end of the HR donor, as the genomic end itself cannot be elongated by a DNA polymerase.

It will be interesting to test whether it will be possible to further improve the efficiency of *in planta* GT by expressing Cas12a (Cpf1) (Zetsche *et al.*, 2015, 2017a,b; Kleinstiver *et al.*, 2016), instead of Cas9, under the control of the egg cell-specific promoter. Cas12a produces short 5′ overhangs and repair of the induced DSB by NHEJ does not often result in mutations of the PAM and sgRNA binding sequence, thus enhancing the opportunity for DSB-induced repair by HR. Although based on small numbers, a recent study using biolistic transformation of rice indicated that with the use of Cas12a from *Francisella novicida* GT of up to 8% can be achieved (Begemann *et al.*, 2017) and by the use of Cas12a RNPs from Lachnospiraceae bacterium *ND2006* up to 10% in *Chlamydomonas* (Ferenczi *et al.*, 2017).

Due to the lack of satisfying frequencies, many attempts have been undertaken using synthetic nucleases to enhance GT frequencies in plants by one or other means [for review see (Steinert *et al.*, 2016)]. However, as it is hard to obtain a bigger number of positive GT events in plants, many studies draw their conclusions on a relatively small sample size, with the danger of overestimating the success of the respective means. In contrast, in this study we produced a large data set obtaining more than 1000 heritable GT events in the T2 generation. We are therefore confident that the use of SaCas9 in combination with the egg cell specific promoter will be an important step forward in giving standard plant molecular biology laboratories the possibility to change Arabidopsis genes in a predefined manner. In this study we used selection to detect targeting events. However, the frequency we obtained was in the 1% range for most lines. This makes it feasible to identify perfect GT events with simple molecular screening protocols, without the need of any selection. A simple PCR screen of T2 seedlings of a few 100 lines, at most, transformed with a single T-DNA containing the HR donor and the Cas9 expression cassette should guarantee the identification of the line with the desired genome modification.

EXPERIMENTAL PROCEDURES

T-DNA constructs used in this study

We used the previously described Gateway® (Thermo Fisher Scientific Inc., https://www.thermofisher.com) compatible pDe-Sa-Cas9 and pEn-Sa-Chimera plasmids (Steinert *et al.*, 2016) for the *S. aureus* based constructs. For the *S. pyogenes* Cas9 nickase-based constructs, we used the previously described pEn-C1.1 and pDe-Cas9-D10A (Schiml *et al.*, 2014), although the kanamycin resistance cassette was exchanged to gentamycin via *Hin*dIII, to allow co-transformation with pDe-Sa-Cas9 based constructs, which also carry a kanamycin resistance gene. The SaCas9 nickase was constructed by PCR-based site-directed mutagenesis. CRISPR target sequences were inserted as annealed oligonucleotides into pEn-Sa-Chimera and pEn-C1.1 using FW1–FW10 (see Table S2 for oligos used in the study). The programmed sgRNA expression cassettes were transferred from the entry vectors into pDe-Sa-Cas9 or pDe-Cas9-D10A-Gent via the Gateway LR reaction. When two sgRNAs were inserted into the destination vectors, the first sgRNA was transferred via the *Mlu*I and *Bsu*36I sites, and the second sgRNA was transferred via LR reaction as previously described (Schiml *et al.*, 2014). The HR donor was synthesized by GeneArt® (Thermo Fisher Scientific Inc.) and consisted of the following features in the given order: *Aat*II site, protospacer and PAM (NNGGGT) for SaCas9, upstream homology, silent mutations at the cutting site, downstream homology containing the S653N mutation, protospacer and PAM (NNGGGT) for SaCas9, *Pac*I site. The HR donor was inserted into the sgRNA containing destination vectors in reverse orientation using *Aat*II and *Pac*I sites (for sequence information of final constructs, see Sequences S1–S10). The CLAVATA3 and YAO promoters were amplified from Arabidopsis genomic DNA using primers FW18 + FW19 or JP20 + JP21, respectively, and inserted into pDe-Sa-Cas9 in the place of the PcUbiquitin4-2 promoter, via *Eco*RI. The egg cell-specific promoter was amplified from pHEE2E-TRI (Wang *et al.*, 2015) (ordered from Addgene), using FW37 + FW38, and inserted into pDe-Sa-Cas9 analogously. The CLAVATA3 terminator was amplified from Arabidopsis genomic DNA using FW20 + FW55 and inserted into pDe-Sa-Cas9-CLVP in the place of the pea3A-terminator, via *Sac*I. The rbcS-E9 terminator was amplified from PHEE2E-TRI using primers FW39 + FW40, and inserted into pDe-Sa-Cas9-ECP analogously. sgRNA cassettes were inserted into the destination vectors with exchanged promoters and/or terminators as described above. The silent mutations on the HR donor were adjusted for the 3′ overhang approach via Gibson Assembly® (New England Biolabs, NEB, https://www.neb.com/). The silent mutations on the resulting HR donor were further modified for the 5′ overhang approach via PCR-based site-directed-mutagenesis (following the NEB protocol for Q5 polymerase) using primers FW66 + FW67. All plasmids are freely available on request.

Plant transformation, selection and handling

All Arabidopsis plants used in this study were of the Columbia-0 background. They were transformed via *Agrobacterium*-mediated transformation as previously described (Clough and Bent, 1998; *Agrobacterium* strain GV3101). Seeds were sown on agar plates containing germination medium [GM: 4.9 g/L Murashige & Skoog medium (Murashige and Skoog, 1962), 10 g/L saccharose, pH 5.7, 7.6 g/L plant-agar] or on a substrate containing 1:1 Floraton 3 (Floragard Vertriebs GmbH, www.floragard.de, Oldenburg, Germany) and Vermiculite (Deutsche Vermiculite Dämmstoff GmbH, www.vermiculite.de, Sprockhövel, Germany).

In pilot experiments, we tested different concentrations of imazapyr (IM) for the efficient identification of GT events. This was done by generating lines with a T-DNA transgene with an ALS ORF carrying the changes that would also be introduced by successful GT. Thus, we defined the optimal IM concentration for selection as 5 μM. Primary (co)transformants carrying the GT construct(s) were first identified under axenic conditions on the respective selection media (kanamycin and/or gentamycin) and then transferred to soil substrate and grown to maturity in the green house. Seeds were harvested for each plant separately. The total weight was determined to estimate seed numbers. After surface sterilization using 4% sodium hypochlorite, and stratification at 4°C overnight, the seeds for each line were sown independently on 5 μM IM-containing medium. After 14 days of growth, the number of resistant seedlings were counted. DNA of representative seedlings was purified for molecular analysis as described previously (Edwards *et al.*, 1991).

Analysis of β-glucuronidase activity

For the intrachromosomal recombination assay, transgenic seedlings carrying transgenes of both the recombination trap and the respective Cas9 expression cassette were harvested 14 days after sowing on agar plates containing germination medium (GM) with corresponding selection markers and cefotaxime. Extraction and quantification of β-glucuronides activity by a 4-MUG assay was performed as previously described (Fauser *et al.*, 2014).

Molecular characterization of GT events

To characterize precise GT events, amplicons were generated from purified DNA of IM-resistant seedlings using primers FW58/FW183, followed by Sanger sequencing (GATC Biotech) using primers FW56 to check for S653N, silent mutations and the downstream junction and FW42 to check the upstream junction. To detect ectopic GT events, amplicons were generated using primers FW57/FW183 and Sanger sequencing was performed as described above using FW56 and FW42.

ACKNOWLEDGEMENTS

We thank Kristina Chajdova, Luisa Telpl and Waltraud Wehrle for excellent technical assistance. This work was funded by the German Federal Ministry of Education and Research (PLANT 2030 – HaploTools, Grant 031B0192A).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Model for ectopic GT.

Figure S2. ALS mutagenesis in plants with ectopic GT events.

Figure S3. Design of the silent mutations on the HR donors.

Table S1. Detailed representation of the results of the molecular analysis of all IM-resistant plants in relation to their respective T1 lines.

Table S2. Oligos used in this study.

Sequence S1. GT vector.

Sequence S2. SaCas9 expression cassette with ubiquitin promoter.

Sequence S3. sgRNA expression system for *S. aureus*.

Sequence S4. SpCas9-D10A expression cassette with ubiquitin promoter.

Sequence S5. sgRNA expression system for *S. pyogenes*.

Sequence S6. EC1 promoter sequence used in this study.

Sequence S7. rbcS-E9 terminator sequence used in this study.

Sequence S8. CLAVATA3 promoter sequence used in this study.

Sequence S9. CLAVATA 3 terminator sequence used in this study.

Sequence S10. YAO promoter sequence used in this study.

REFERENCES

Altpeter, F., Springer, N.M., Bartley, L.E. *et al.* (2016) Advancing crop transformation in the era of genome editing. *Plant Cell*, **28**, 1510–1520.

Ayar, A., Wehrkamp-Richter, S., Laffaire, J.-B. *et al.* (2013) Gene targeting in maize by somatic ectopic recombination. *Plant Biotechnol. J.* **11**, 305–314.

Baltes, N.J., Gil-Humanes, J., Čermák, T., Atkins, P.A. and Voytas, D.F. (2014) DNA replicons for plant genome engineering. *Plant Cell*, **26**, 151–163.

Baltes, N.J., Gil-Humanes, J. and Voytas, D.F. (2017) Genome engineering and agriculture. Opportunities and challenges. *Prog. Mol. Biol. Transl. Sci.* **149**, 1–26.

Begemann, M.B., Gray, B.N., January, E., Gordon, G.C., He, Y., Liu, H., Wu, X., Brutnell, T.P., Mockler, T.C. and Oufattole, M. (2017) Precise insertion and guided editing of higher plant genomes using Cpf1 CRISPR nucleases. *Sci. Rep.* **7**, 11606.

Bothmer, A., Phadke, T., Barrera, L.A. *et al.* (2017) Characterization of the interplay between DNA repair and CRISPR/Cas9-induced DNA lesions at an endogenous locus. *Nat. Commun.* **8**, 13905.

Butt, H., Eid, A., Ali, Z., Atia, M.A.M., Mokhtar, M.M., Hassan, N., Lee, C.M., Bao, G. and Mahfouz, M.M. (2017) Efficient CRISPR/Cas9-mediated genome editing using a chimeric single-guide RNA molecule. *Front. Plant Sci.* **8**, 1441.

Čermák, T., Baltes, N.J., Čegan, R., Zhang, Y. and Voytas, D.F. (2015) High-frequency, precise modification of the tomato genome. *Genome Biol.* **16**, 232.

Čermák, T., Curtin, S.J., Gil-Humanes, J. *et al.* (2017) A multipurpose toolkit to enable advanced genome engineering in plants. *Plant Cell*, **29**, 1196–1217.

Clough, S.J. and Bent, A.F. (1998) Floral dip. A simplified method forAgrobacterium-mediated transformation ofArabidopsis thaliana. *Plant J.* **16**, 735–743.

Daley, J.M., Niu, H., Miller, A.S. and Sung, P. (2015) Biochemical mechanism of DSB end resection and its regulation. *DNA Repair*, **32**, 66–74.

Davis, L. and Maizels, N. (2014) Homology-directed repair of DNA nicks via pathways distinct from canonical double-strand break repair. *Proc. Natl Acad. Sci. USA*, **111**, E924–E932.

D'Halluin, K., Vanderstraeten, C., Stals, E., Cornelissen, M. and Ruiter, R. (2008) Homologous recombination. A basis for targeted genome optimization in crop species such as maize. *Plant Biotechnol. J.* **6**, 93–102.

Edwards, K., Johnstone, C. and Thompson, C. (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* **19**, 1349.

Endo, M., Osakabe, K., Ichikawa, H. and Toki, S. (2006) Molecular characterization of true and ectopic gene targeting events at the acetolactate synthase gene in Arabidopsis. *Plant Cell Physiol.* **47**, 372–379.

Endo, M., Mikami, M. and Toki, S. (2016) Biallelic gene targeting in rice. *Plant Physiol.* **170**, 667–677.

van Ex, F., Verweire, D., Claeys, M., Depicker, A. and Angenon, G. (2009) Evaluation of seven promoters to achieve germline directed Cre-lox recombination in *Arabidopsis thaliana*. *Plant Cell Rep.* **28**, 1509–1520.

Fauser, F., Roth, N., Pacher, M., Ilg, G., Sánchez-Fernández, R., Biesgen, C. and Puchta, H. (2012) In planta gene targeting. *Proc. Natl Acad. Sci. USA*, **109**, 7535–7540.

Fauser, F., Schiml, S. and Puchta, H. (2014) Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *Plant J.* **79**, 348–359.

Ferenczi, A., Pyott, D.E., Xipnitou, A. and Molnar, A. (2017) Efficient targeted DNA editing and replacement in *Chlamydomonas reinhardtii* using

Cpf1 ribonucleoproteins and single-stranded DNA. *Proc. Natl Acad. Sci. USA*, **114**, 13567–13572.

Gil-Humanes, J., Wang, Y., Liang, Z. *et al.* (2017) High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. *Plant J.* **89**, 1251–1262.

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.

Kleinstiver, B.P., Tsai, S.Q., Prew, M.S., Nguyen, N.T., Welch, M.M., Lopez, J.M., McCaw, Z.R., Aryee, M.J. and Joung, J.K. (2016) Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat. Biotechnol.* **34**, 869–874.

Knoll, A., Fauser, F. and Puchta, H. (2014) DNA recombination in somatic plant cells: mechanisms and evolutionary consequences. *Chromosome Res.* **22**, 191–201.

Lee, K.Y., Townsend, J., Tepperman, J., Black, M., Chui, C.F., Mazur, B., Dunsmuir, P. and Bedbrook, J. (1988) The molecular basis of sulfonylurea herbicide resistance in tobacco. *EMBO J.* **7**, 1241–1248.

Mahas, A., Neal Stewart, C. and Mahfouz, M.M. (2017) Harnessing CRISPR/Cas systems for programmable transcriptional and post-transcriptional regulation. *Biotechnology Adv.* **36**, 295–310.

Mahfouz, M.M. (2017) Genome editing: the efficient tool CRISPR-Cpf1. *Nat. Plants*, **3**, 17028.

Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L. and Church, G.M. (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* **31**, 833–838.

Malzahn, A., Lowder, L. and Qi, Y. (2017) Plant genome editing with TALEN and CRISPR. *Cell Biosci.* **7**, 21.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* **15**, 473–497.

Nishizawa-Yokoi, A., Čermák, T., Hoshino, T. *et al.* (2016) A defect in DNA ligase4 enhances the frequency of TALEN-mediated targeted mutagenesis in rice. *Plant Physiol.* **170**, 653–666.

Orel, N. and Puchta, H. (2003) Differences in the processing of DNA ends in *Arabidopsis thaliana* and tobacco: possible implications for genome evolution. *Plant Mol. Biol.* **51**, 523–531.

Pacher, M. and Puchta, H. (2017) From classical mutagenesis to nuclease-based breeding - directing natural DNA repair for a natural end-product. *Plant J.* **90**, 819–833.

Paix, A., Folkmann, A., Goldman, D.H., Kulaga, H., Grzelak, M.J., Rasoloson, D., Paidemarry, S., Green, R., Reed, R.R. and Seydoux, G. (2017) Precision genome editing using synthesis-dependent repair of Cas9-induced DNA breaks. *Proc. Natl Acad. Sci. USA*, **114**, E10745–E10754.

de Pater, S., Pinas, J.E., Hooykaas, P.J.J. and van der Zaal, B.J. (2013) ZFN-mediated gene targeting of the Arabidopsis protoporphyrinogen oxidase gene through Agrobacterium-mediated floral dip transformation. *Plant Biotechnol. J.* **11**, 510–515.

Puchta, H. (1998) Repair of genomic double-strand breaks in somatic plant cells by one-sided invasion of homologous sequences. *Plant J.* **13**, 331–339.

Puchta, H. (1999) Double-strand break-induced recombination between ectopic homologous sequences in somatic plant cells. *Genetics*, **152**, 1173–1181.

Puchta, H. (2016) Using CRISPR/Cas in three dimensions: towards synthetic plant genomes, transcriptomes and epigenomes. *Plant J.* **87**, 5–15.

Puchta, H. (2017) Applying CRISPR/Cas for genome engineering in plants. The best is yet to come. *Curr. Opin. Plant Biol.* **36**, 1–8.

Puchta, H. and Fauser, F. (2013) Gene targeting in plants: 25 years later. *Int. J. Dev. Biol.* **57**, 629–637.

Puchta, H. and Hohn, B. (2012) In planta somatic homologous recombination assay revisited: a successful and versatile, but delicate tool. *Plant Cell*, **24**, 4324–4331.

Puchta, H., Dujon, B. and Hohn, B. (1996) Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proc. Natl Acad. Sci. USA*, **93**, 5055–5060.

Roth, N., Klimesch, J., Dukowic-Schulze, S., Pacher, M., Mannuss, A. and Puchta, H. (2012) The requirement for recombination factors differs considerably between different pathways of homologous double-strand break repair in somatic plant cells. *Plant J.* **72**, 781–790.

Salomon, S. and Puchta, H. (1998) Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO J.* **17**, 6086–6095.

Sauer, N.J., Narvaez-Vasquez, J., Mozoruk, J. *et al.* (2016) Oligonucleotide-mediated genome editing provides precision and function to engineered nucleases and antibiotics in plants. *Plant Physiol.* **170**, 1917–1928.

Scheben, A., Wolter, F., Batley, J., Puchta, H. and Edwards, D. (2017) Towards CRISPR/Cas crops - bringing together genomics and genome editing. *New Phytol.* **216**, 682–698.

Schiml, S. and Puchta, H. (2016) Revolutionizing plant biology: multiple ways of genome engineering by CRISPR/Cas. *Plant Methods*, **12**, 8.

Schiml, S., Fauser, F. and Puchta, H. (2014) The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. *Plant J.* **80**, 1139–1150.

Schiml, S., Fauser, F. and Puchta, H. (2016) Repair of adjacent single-strand breaks is often accompanied by the formation of tandem sequence duplications in plant genomes. *Proc. Natl Acad. Sci. USA*, **113**, 7266–7271.

Shukla, V.K., Doyon, Y., Miller, J.C. *et al.* (2009) Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature*, **459**, 437–441.

Sprunck, S., Rademacher, S., Vogler, F., Gheyselinck, J., Grossniklaus, U. and Dresselhaus, T. (2012) Egg cell-secreted EC1 triggers sperm cell activation during double fertilization. *Science*, **338**, 1093–1097.

Steinert, J., Schiml, S., Fauser, F. and Puchta, H. (2015) Highly efficient heritable plant genome engineering using Cas9 orthologues from *Streptococcus thermophilus* and *Staphylococcus aureus*. *Plant J.* **84**, 1295–1305.

Steinert, J., Schiml, S. and Puchta, H. (2016) Homology-based double-strand break-induced genome engineering in plants. *Plant Cell Rep.* **35**, 1429–1438.

Svitashev, S., Young, J.K., Schwartz, C., Gao, H., Falco, S.C. and Cigan, A.M. (2015) Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. *Plant Physiol.* **169**, 931–945.

Townsend, J.A., Wright, D.A., Winfrey, R.J., Fu, F., Maeder, M.L., Joung, J.K. and Voytas, D.F. (2009) High-frequency modification of plant genes using engineered zinc-finger nucleases. *Nature*, **459**, 442–445.

Vriend, L.E.M., Prakash, R., Chen, C.-C., Vanoli, F., Cavallo, F., Zhang, Y., Jasin, M. and Krawczyk, P.M. (2016) Distinct genetic control of homologous recombination repair of Cas9-induced double-strand breaks, nicks and paired nicks. *Nucleic Acids Res.* **44**, 5204–5217.

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.

Wang, M., Lu, Y., Botella, J.R., Mao, Y., Hua, K. and Zhu, J.-K. (2017) Gene targeting by homology-directed repair in rice using a geminivirus-based CRISPR/Cas9 system. *Mol. Plant*, **10**, 1007–1010.

Watanabe, K., Breier, U., Hensel, G., Kumlehn, J., Schubert, I. and Reiss, B. (2016) Stable gene replacement in barley by targeted double-strand break induction. *J. Exp. Bot.* **67**, 1433–1445.

Wright, D.A., Townsend, J.A., Winfrey, R.J.J.R., Irwin, P.A., Rajagopal, J., Lonosky, P.M., Hall, B.D., Jondle, M.D. and Voytas, D.F. (2005) High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant J.* **44**, 693–705.

Yan, L., Wei, S., Wu, Y., Hu, R., Li, H., Yang, W. and Xie, Q. (2015) High-efficiency genome editing in Arabidopsis using YAO promoter-driven CRISPR/Cas9 system. *Mol. Plant*, **8**, 1820–1823.

Yin, K., Gao, C. and Qiu, J.-L. (2017) Progress and prospects in plant genome editing. *Nat. Plants*, **3**, 17107.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O. *et al.* (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, **163**, 759–771.

Zetsche, B., Heidenreich, M., Mohanraju, P. *et al.* (2017a) Erratum: Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat. Biotechnol.* **35**, 178.

Zetsche, B., Heidenreich, M., Mohanraju, P. *et al.* (2017b) Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat. Biotechnol.* **35**, 31–34.