the plant journal

SEB Society for Experimental Biology

The Plant Journal (2019) doi: 10.1111/tpj.14322

TECHNICAL ADVANCE

Efficient induction of heritable inversions in plant genomes using the CRISPR/Cas system

Carla Schmidt, Michael Pacher and Holger Puchta*

Botanical Institute, Fritz-Haber-Weg 4, 76133 Karlsruhe, Germany

Received 6 February 2019; revised 11 March 2019; accepted 15 March 2019.
*For correspondence (e-mail holger.puchta@kit.edu).

SUMMARY

During the evolution of plant genomes, sequence inversions occurred repeatedly making the respective regions inaccessible for meiotic recombination and thus for breeding. Therefore, it is important to develop technologies that allow the induction of inversions within chromosomes in a directed and efficient manner. Using the Cas9 nuclease from *Staphylococcus aureus* (*Sa*Cas9), we were able to obtain scarless heritable inversions with high efficiency in the model plant *Arabidopsis thaliana*. Via deep sequencing, we defined the patterns of junction formation in wild-type and in the non-homologous end-joining (NHEJ) mutant *ku70-1*. Surprisingly, in plants deficient of KU70, inversion induction is enhanced, indicating that KU70 is required for tethering the local broken ends together during repair. However, in contrast to wild-type, most junctions are formed by microhomology-mediated NHEJ and thus are imperfect with mainly deletions, making this approach unsuitable for practical applications. Using egg-cell-specific expression of Cas9, we were able to induce heritable inversions at different genomic loci and at intervals between 3 and 18 kb, in the percentage range, in the T1 generation. By screening individual lines, inversion frequencies of up to the 10% range were found in T2. Most of these inversions had scarless junctions and were without any sequence change within the inverted region, making the technology attractive for use in crop plants. Applying our approach, it should be possible to reverse natural inversions and induce artificial ones to break or fix linkages between traits at will.

Keywords: chromosomal rearrangements, inversions, genome engineering, double-strand break repair, *Arabidopsis thaliana*, Cas9, technical advance.

INTRODUCTION

The targeted induction of site-specific double-strand breaks (DSBs) is the basis of editing genes in eukaryotes, including plants. As the repair of such breaks occurs either by non-homologous end joining (NHEJ) or homologous recombination (HR; Puchta, 2005), different genomic changes such as the knock-out of genes (Salomon and Puchta, 1998) or the knock-in of specific sequences (Puchta *et al.*, 1996) can be achieved. Although a number of programmable site-specific nucleases have been developed over time (Voytas, 2013; Puchta and Fauser, 2014), the field of plant genome engineering was revolutionized with the elucidation of the properties of the Cas9 nuclease (Puchta, 2017; Gao, 2018). With the application of multiple sgRNAs, the CRISPR/Cas system can be used to induce multiple DSBs at once, therefore enabling for complex chromosomal modifications to be achieved (Jinek *et al.*, 2012; Le Cong *et al.*, 2013).

By inducing two DSBs, a greater step from gene editing to genome engineering can be taken. If the two breaks are located on different plant chromosomes, reciprocal translocations can be obtained (Pacher *et al.*, 2007), whereas the induction of two breaks on the same chromosome mainly leads to deletions (Siebert and Puchta, 2002) and sometimes inversions (Qi *et al.*, 2013a; Zhang *et al.*, 2017). Indeed, the CRISPR/Cas system has been successfully used in mammals for the induction of deletions, inversions and translocations (for review, see Cheong *et al.*, 2018). In plants, on the other hand, most reports focused on the induction of deletions and their inheritance into the next generation. Deletions were induced in *Arabidopsis thaliana* using different promoters for stage- or tissue-specific expression of Cas9, obtaining a deletion size of up to 13 kb with high frequencies of heritable events or up to 120 kb with low frequencies (Ordon *et al.*, 2017; Durr *et al.*, 2018;

Wu *et al.*, 2018). For stage-specific expression, the egg cell promoter is commonly used for early DSB induction to create heritable mutations, large deletions and gene-targeting events in Arabidopsis (Wang *et al.*, 2015; Durr *et al.*, 2018; Wolter *et al.*, 2018). Moreover, chromosomal deletions up to 58 kb were induced using Cas9 in *Medicago truncatula*, and up to 3 kb in tomato protoplasts. The deletions in *M. truncatula* could also be successfully transferred into the germ line (Čermák *et al.*, 2017).

Genome rearrangements occur regularly in plants (Udall *et al.*, 2005; Szinay *et al.*, 2012; Li *et al.*, 2016; Zapata *et al.*, 2016), in particular inversions are associated across species mainly with adaptation, specification and genome evolution (Blanc *et al.*, 2000; Navarro and Barton, 2003; Kirkpatrick and Barton, 2006; Fang *et al.*, 2012; Schubert and Vu, 2016). In natural populations, the most common large-scale chromosomal structural variations are inversions, with changes leading to hybrid sterility, centromere shifting, the formation of new open reading frames (ORFs), disruption of already existing genes, alteration of expression profiles, and also the formation or breakage of genetic linkages (Madan, 1995; Lowry and Willis, 2010; Schubert, 2018). Probably the most popular paracentric inversion in *A. thaliana* is the heterochromatic knob hk4S with the associated centromere shift on the short arm of chromosome four, which was identified in the Columbia accession but is absent in the Landsberg accession (Fransz *et al.*, 2016). Moreover, there are inversions in tomato, leading to hindrances of introgressions of resistance markers from wild relatives to cultivated tomato species and to the separation of promoter regions from their ORFs, resulting in a moderate effect on locule number (Lin *et al.*, 2014; Rodríguez-Leal *et al.*, 2017).

Inversions principally act as major obstacles for breeding as no crossovers can be achieved between inverted regions and trait linkages cannot be broken. On the other hand, it might also be advantageous for breeders to stabilize genetic linkages in elite cultivars, which should be achievable by the induction of programmable site-specific inversions. Therefore, it is highly desirable to set up a technology in plants for the efficient induction of heritable inversions in a sequence-independent manner. Here we develop such a technology for the model plant *A. thaliana*. Targeted inversions at four different loci on three different chromosomes with a size up to 18 kb were achieved. Most inversions were scarless with perfect ligation of all broken ends, and no mutations were induced within the inverted sequence. Using the egg cell promoter, we were able to obtain heritable inversions in up to 10% of the progeny, depending on the transgenic line.

RESULTS

Induction of inversions in the Arabidopsis genome

To obtain knowledge of how frequently inversions may occur in *A. thaliana* when two DSBs are induced in close proximity on the same chromosome, we chose four different loci on three different chromosomes, namely the genes *ADH1* (AT1G77120, alcohol dehydrogenase 1), *TT4* (AT5G13930, transparent testa 4) and *ECA3* (AT1G10130, ER-type Ca^{2+}-ATPase 3), as well as a pseudogene (PSDG, AT2G05640). The two DSBs leading to the inversion were induced with a distance of about 3 kb between them, at all four loci (Figure 1a). While the protospacer of the *ECA3* approach was located within the ORF of the gene, the protospacers of the *TT4* and *ADH1* approaches were located outside the ORF. The appropriate spacer sequences were cloned into the pDe-*Sa*-Cas9 vector, where *Sa*Cas9 is under the control of the *Pc*Ubi4-2 promoter for constitutive expression (Steinert *et al.*, 2015, 2017). Following *Agrobacterium tumefaciens*-mediated floral-dip transformation, the *A. thaliana* seedlings were grown on selection media for 2 weeks to identify stable transformants. We decided to measure the efficiency of inversion formation by determining the amount of both of the newly formed junctions, which we termed proximal (p) or distal (d), in accordance with their orientation to the centromere. We analyzed three biological replicates, with the extracted DNA of 100 independent primary transformants for each, and quantified the amount of inversions by digital droplet polymerase chain reaction (ddPCR) using site-specific primers (Figure 1b). The number of detected events was set in relation to the number of genomes by measuring the amplification of a genomic fragment on the same chromosome in close proximity, outside of the inverted region. We were able to detect inversions at all four loci in 0.5–2% of the examined genomes. The amount of p and d junctions was highly similar for their respective loci, indicating that the analysis was reliable and reproducible (Figure 1c). As the induction of two DSBs in close proximity can also lead to the deletion of the intervening sequence, we also measured the amount of deletions via ddPCR. Deletions occurred in all four loci with higher frequencies than inversions, between 1 and 7% (Figure 1d). There is a clear indication that the efficiency of inversion and deletion formation is influenced by the locus, hinting that the efficiency of DSB induction might be rate limiting, as it is the prerequisite for both kinds of rearrangements. This was confirmed by determining the cutting efficiency at both break sites for all four loci using a T7 endonuclease assay (Figure S1). Indeed, DSB induction was most efficient at the *PSDG* locus, which showed higher inversion and deletion frequencies compared with the other three loci.

To characterize the mechanism of inversion formation, we amplified both the p and d junctions of the inversions at the *ADH1* and *TT4* loci, and performed deep sequencing with these amplicons. Interestingly, the majority of reads showed a precise ligation without any sequence change, with rejoined breakpoints 3 bp upstream of the respective PAM sequences (Figure 1e). For the *ADH1* locus, in about

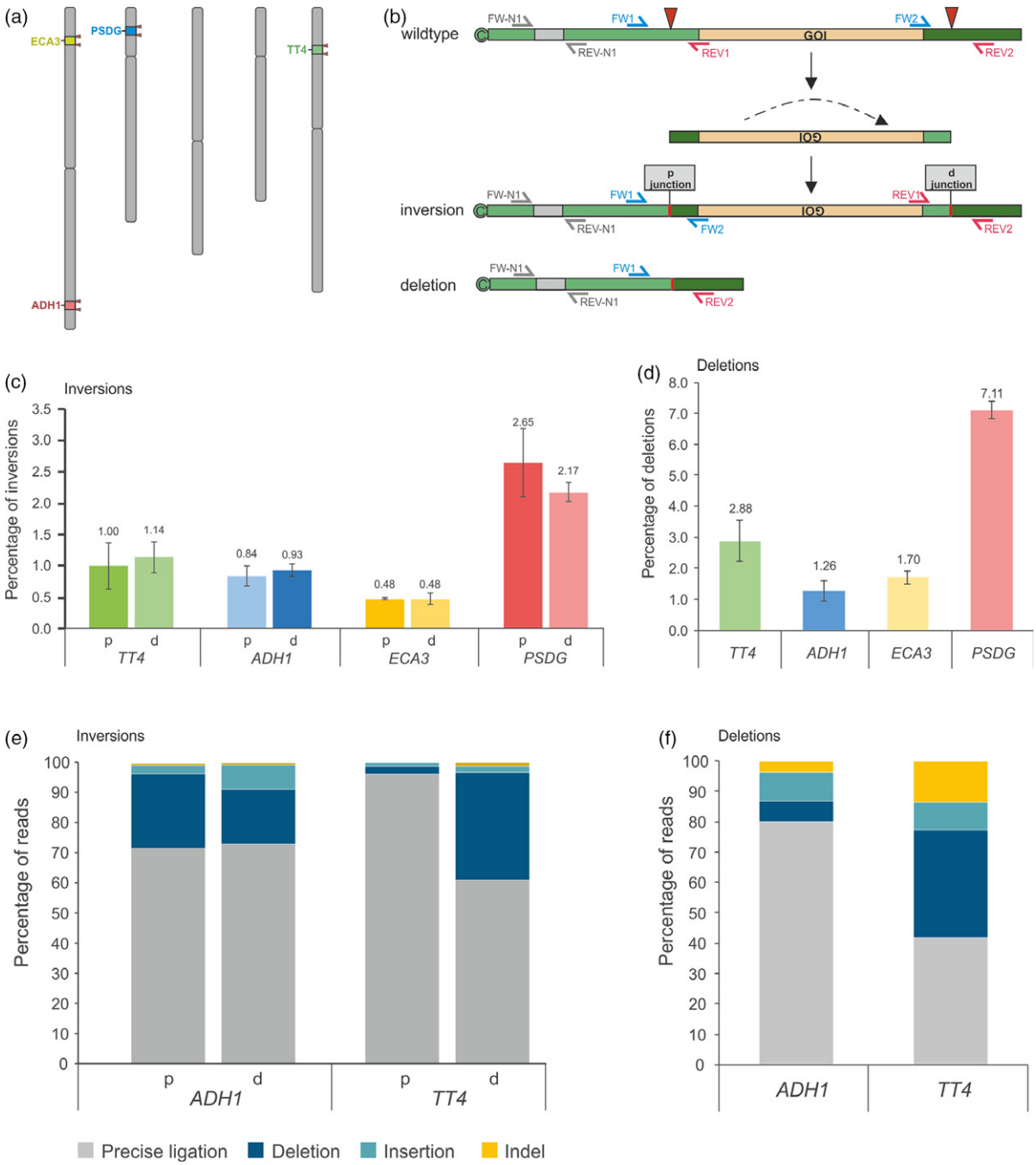


Figure 1. Quantitative and qualitative analysis of *Sa*Cas9-induced inversion and deletion events in wild-type.
(a) Inversions and deletions were induced at four different loci on three different chromosomes in the Arabidopsis genome.
(b) Schematic representation of the formation of an inversion event after induction of two double-strand breaks (DSBs; red triangles), surrounding a gene of interest (GOI, yellow). After break induction, the fragment can be released and incorporated in an inverted orientation. The inversion can be detected using site-specific primers for both newly formed junctions. Primers FW1 and FW2 (blue) are specific for the p junction, and primers REV1 and REV2 (red) are specific for the d junction. Deletion events can be detected by combining the primers FW1 and REV2. Additionally, a third primer combination (gray), located upstream of the inverted region, was used as a control to set the value in relation to the absolute number of genomes analyzed.
The amount of inversions (c) and deletions (d) relative to the genome number as percentage were determined via digital droplet polymerase chain reaction (ddPCR) in wild-type background. For all four approaches, the amount of both p and d junctions of the inversion was measured, showing no significant difference, and additionally the amount of deletion events was determined. Each data point consists of three independent biological replicates and each replicate consisted of 100 independent T1 plants ($n = 3$). Error bars represent the standard deviation of the replicates.
Deep sequencing results of the inversion (e) and deletion (f) junctions of the *ADH1* and the *TT4* locus. Reads were grouped in four different classes related to their mutations. The most abundant class of reads showed precise ligation events for all four junctions of the inversions and also for deletion junctions of the *ADH1* locus. At the *TT4* locus about 40% of the analyzed reads were precisely ligated.

three-quarters of the cases of both p and d junction's, simple re-ligation occurred. In comparison, for *TT4*, almost all p junctions arose via simple re-ligation, as this was only the case for less than two-thirds of the d junctions and this difference correlates with the presence of putative microhomologies (MHs) near the d junction. In case of the p junction, there is an underrepresentation of putative MHs. Taking only the reads into account that show some deletions, this difference becomes even clearer. Here, half of all reads representing the p junctions were formed without the use of MHs. In contrast, three-quarters of the reads at the d junction harboring a deletion were formed by the use of a specific 3-bp MH (Figure S2). Nevertheless, imperfect junctions show only a loss of one or few nucleotides. Insertions or combinations of both a deletion and an insertion (indel) were only found to a limited extent. The same analysis was performed for the composition of deletions, and a similar distribution of sequence patterns could be found as for inversions (Figure 1f). This strongly indicates that for both kinds of changes, the rejoining of the broken ends occurs by identical mechanisms.

KU70 limits inversion formation and is involved in the formation of scarless junctions

Due to the fact that most inversions had junctions devoid of deletions and other mutations, we assumed that the classical NHEJ pathway with the KU heterodimer as an end-protecting factor was required for inversion formation. Therefore, we decided to use the well-characterized DNA repair mutant *ku70-1*, which is deficient in the cNHEJ pathway (Tamura *et al.*, 2002). *Ku70-1* mutant plants were transformed with the *ADH1* and the *TT4* constructs, and the amount of inversions, as well as the sequence composition of the respective junctions, was determined. Surprisingly, the results of the ddPCR showed that inversion and deletion frequencies were enhanced in the *ku70-1* mutant compared with wild-type (Figure 2a). In the *ADH1* locus, about a twofold increase in the formation of both p and d junctions was found, in comparison to wild-type. The same trend could be observed for both junctions at the *TT4* locus. We were completely surprised by this finding, as it demonstrated that KU70 is indeed suppressing inversion formation. This indicated that besides its end-binding function, the KU heterodimer is also required for tethering the broken ends together. In its absence, other pathways are able to take over NHEJ efficiently. However, this comes with a price for genome stability, as formerly unlinked broken ends are now joined with higher efficiency. If our hypothesis was correct, we should see that deletion formation is enhanced in the mutant background, as well. Indeed, medium values of deletions obtained from both loci were higher in the mutant background; in case of *TT4* we detected a highly significant duplication in the number of deletion events (Figure 2b). According to our expectations, the results of the amplicon deep sequencing demonstrated a very different way of junction formation in the *ku70-1* mutant (Figure 2c). Only a tiny fraction at all four analyzed junctions was formed by precise ligation. Most of the reads showed mutations at the break sites with most being deletions and indels. In contrast, the amount of insertions at junctions obtained with the *ku70-1* mutant was comparable to wild-type or even decreased. Additionally, we detected a significant change in the quality of mutations at the junctions (Figure 2e,f). We observed in wild-type that the length of most deletions and insertions ranged from 1 to 10 bp, and only a small number of reads showed deletions larger than 50 bp. This situation changed completely in the *ku70-1* mutant where most of the deletions were either 11–50 bp long or even longer than 50 bp. For insertion formation, we detected a similar increase for most junctions, with many insertions ranging from 11 to 50 bp in the *ku70-1* mutant background.

In the absence of classical non-homologous end-joining, micro-homology-mediated end-joining can efficiently be used for inversion and deletion formation

To further characterize the underlying repair pathway that was responsible for junction formation in the *ku70-1* mutant background, we checked if MHs played a role in junction formation. It could be demonstrated in plants that mutants lacking KU use a MH-based repair pathway for DSB repair (Qi *et al.*, 2013b; Shen *et al.*, 2017). The more interesting question is whether this MH-based repair pathway is also involved in the formation of inversions. Therefore, we used the 100 most prominent different alleles of the *ADH1* locus and analyzed them in detail with how often we could document the use of MH for joining both ends (Figure 3a,b). Indeed, we found MH use in three-quarters of the cases in the *ku70-1* mutant, whereas such events represented only a minor fraction in wild-type, with less than 7%. Moreover, we obtained large deletions in *ku70-1* with adjacent MHs (Figure 3c) and a large amount of the events repaired via MH were accompanied by indels. Altogether, three-quarters of all analyzed indels of the *ADH1* p junction showed insertions that had their origin directly next to the corresponding break site, with an adjacent MH. In addition, we performed the same analysis for the *TT4* junctions that showed similar results, besides the *ku70-1 TT4* p junction showing only a small number of reads repaired using MHs (Figure S3a,b). This phenomenon is possibly due to a lack of MHs near the break site of the *TT4* p junction. Thus, the loss of KU70 results in an increase of inversion formation by a mutagenic MH-mediated end-joining pathway. As a control, we also performed the same analysis for characterizing the pathway of deletion formation by DSB induction at both loci. In the majority of cases, we also found deletions spanning the DSB sites in *ku70-1* and only a small number of

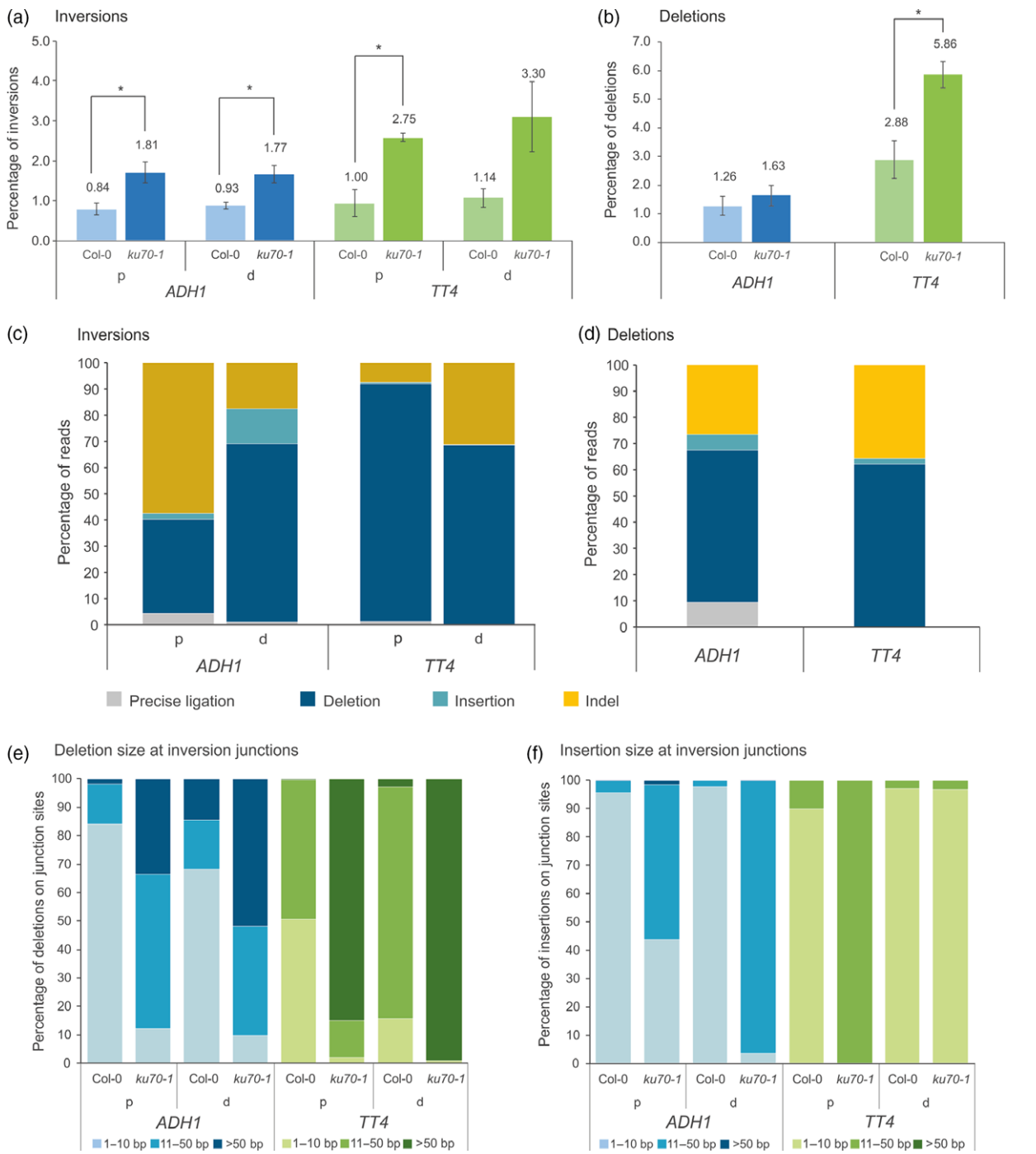
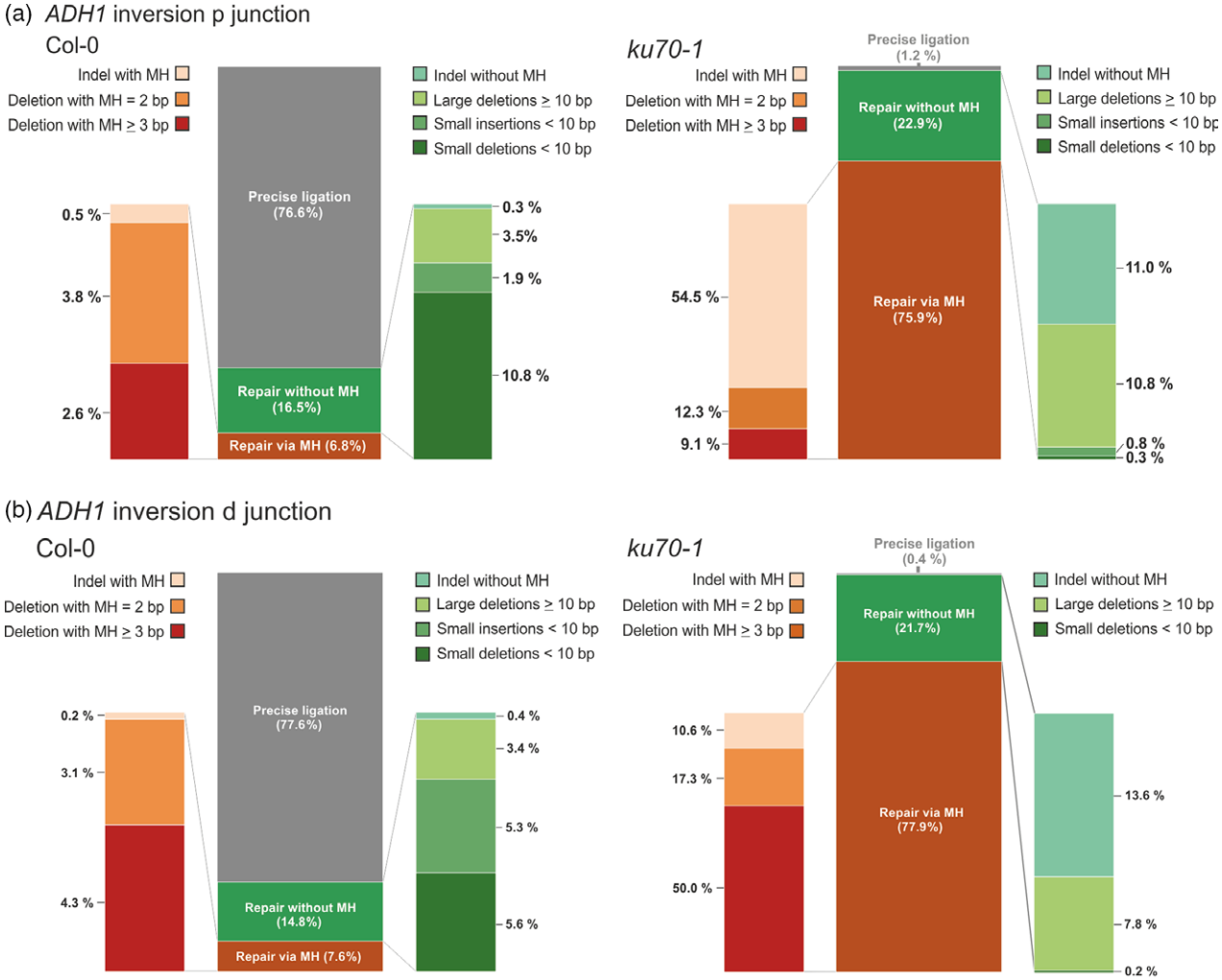


Figure 2. Quantitative and qualitative analysis of *Sa*Cas9-induced inversion and deletion events in the *ku70-1* mutant.
Quantification of both inversion junctions (a) and the deletion junction (b) via digital droplet polymerase chain reaction (ddPCR) in *ADH1* (blue) and *TT4* (green). Increased inversion frequencies were obtained in the *ku70-1* mutant background. Each data point consists of three independent biological replicates and each replicate consisted of 100 independent T1 plants ($n = 3$). Error bars represent the standard deviation of the replicates.
Results of the amplicon deep sequencing of the inversion (c) and the deletion (d) in *ku70-1* background, in which the reads were grouped in four different classes related to their mutation type at the ligation site. Most of the reads showed deletions and indels at ligation sites and only a minor part of the reads showed precise ligated junctions.
(e) Reads were classified according to their deletion size in the *ADH1* (blue) and *TT4* (green) locus. The majority of junctions in wild-type samples showed only small deletions of 1–10 bp in both approaches. In contrast, the deletion size in *ku70-1* mutant samples was much larger, with the majority of junctions showing deletions up to 50 bp and in a higher amount even exceeding 50 bp.
(f) Diagram showing the size of obtained insertions in the *ADH1* and *TT4* loci. In wild-type, most of the analyzed reads contained only small insertions of a few bp and in the *ku70-1* mutant junctions showed large insertions up to 50 bp, while insertions larger than 50 bp were only rarely obtained in all samples.



(c)

Col-0

```
Precise 5' ATTACGAGCAAAATAATTATACAAAAATACTGATAATTATCGAAAGGGTAGTACGAATTCCGCAGGATCTTGATCGGTATAAGGATAGAC 3'
Del1       ATTACGAGCAAAATAATTATACAAAAATACTGATAATTA--GAAAGGGTAGTACGAATTCCGCAGGATCTTGATCGGTATAAGGATAGAC      Δ2
Del2       ATTACGAGCAAAATAATTATACAAAAATACTGATAATTAT---AAGGGTAGTACGAATTCCGCAGGATCTTGATCGGTATAAGGATAGAC      Δ3
Del3       ATTACGAGCAAAATAATTATACAAAAATACTGATAATTAT-GAAAGGGTAGTACGAATTCCGCAGGATCTTGATCGGTATAAGGATAGAC      Δ1
Del4       ATTACGAGCAAAATAATTATACAAAAATACTGATAATTAT----AGGGTAGTACGAATTCCGCAGGATCTTGATCGGTATAAGGATAGAC      Δ4
Del5       ATTACGAGCAAAATAATTATACAAAAATACTGATA-------AAAGGGTAGTACGAATTCCGCAGGATCTTGATCGGTATAAGGATAGAC      Δ7
Del6       ATTACGAGCAAAATAATTATACAAAAATACTGATAATT----AAAGGGTAGTACGAATTCCGCAGGATCTTGATCGGTATAAGGATAGAC      Δ4
Del7       ATTACGAGCAAAATAATTATACAAAAATACTGATAATT------AGGGTAGTACGAATTCCGCAGGATCTTGATCGGTATAAGGATAGAC      Δ6
```

ku70-1

```
Precise 5' ATTACGAGCAAAATAATTATACAAAAATACTGATAATTATCGAAAGGGTAGTACGAATTCCGCAGGATCTTGATCGGTATAAGGATAGAC 3'
Del1       ATTA----------------------------------------------------------CGCAGGATCTTGATCGGTATAAGGATAGAC      Δ55  MH = 2
Del2       ATTACGAGCAAAATAATTATACAAAAATAC-------------------------------------TGATCGGTATAAGGATAGAC      Δ40  MH = 4
Del3       ATTACGAGCAAAATAATTATACAAAAATACTG-----------------------------------------ATAAGGATAGAC      Δ46  MH = 4
Del4       AT--------------------------------------------------------------------TAAGGATAGAC      Δ77  MH = 2
Del5       ATTACGAGCAAAATAATTATACAAAAATACTGATAATTA-----------------------------TTGATCGGTATAAGGATAGAC      Δ33
Del6       ---------- Δ164 nt ----------GATCTTGATCGGTATAAGGATAGAC      Δ164 MH = 2
Del7       ATTACGAGCAAAATAATTATACAAAAATACTGAT-------------------AATTCCGCAGGATCTTGATCGGTATAAGGATAGAC      Δ21  MH = 4
```

Figure 3. Detailed analysis of repair patterns at both inversion junctions of the *ADH1* gene.
Quantification of the occurrence of micro-homologies (MHs) used for junction formation of the *ADH1* p junction (a) and the *ADH1* d junction (b). We distinguished between precisely ligated junctions without alterations (gray), junctions repaired without the use of MHs (green) and repair events related to MH-mediated repair (orange). The different mutations were furthermore grouped within these classes due to their type (deletions, insertions, indels), size and MH size. In wild-type, most junctions showed precise ligation, the second-most frequent reads showed ligation site repair patterns without the use of MHs, and only a minority of junctions were repaired using MHs. Overall in wild-type, the most common mutations obtained were small deletions up to 10 bp. In the *ku70-1* mutant most of the junctions were repaired using MHs, and only a minority of the junctions were precisely ligated or repaired without the use of MHs. In contrast to wild-type, most of the junctions showed large deletions and indels formed via MHs.
(c) Detailed illustration of exemplary junctions with deletions on the sequence level for wild-type and the *ku70-1* mutant. Whereas in wild-type deletions are small, in *ku70-1* large deletions are much more regular with the occurrence of adjacent MHs that were used for junction formation.

junctions were precise (Figure 2d). At both loci in the majority of cases, MHs were involved in the deletion formation in the *ku70-1* mutant background (Figure 4a,b). Thus, in the absence of the classical NHEJ pathway, inversion and deletion formation is due to the MH-mediated NHEJ pathway.

The egg-cell-specific promoter is efficient in inducing inversions that are transferred to the next generation

Our previous analysis showed that although the application of the *ku70-1* mutant background led to higher inversion frequencies, almost all recombinants harbor mutations at both junctions. Thus, the use of a mutant

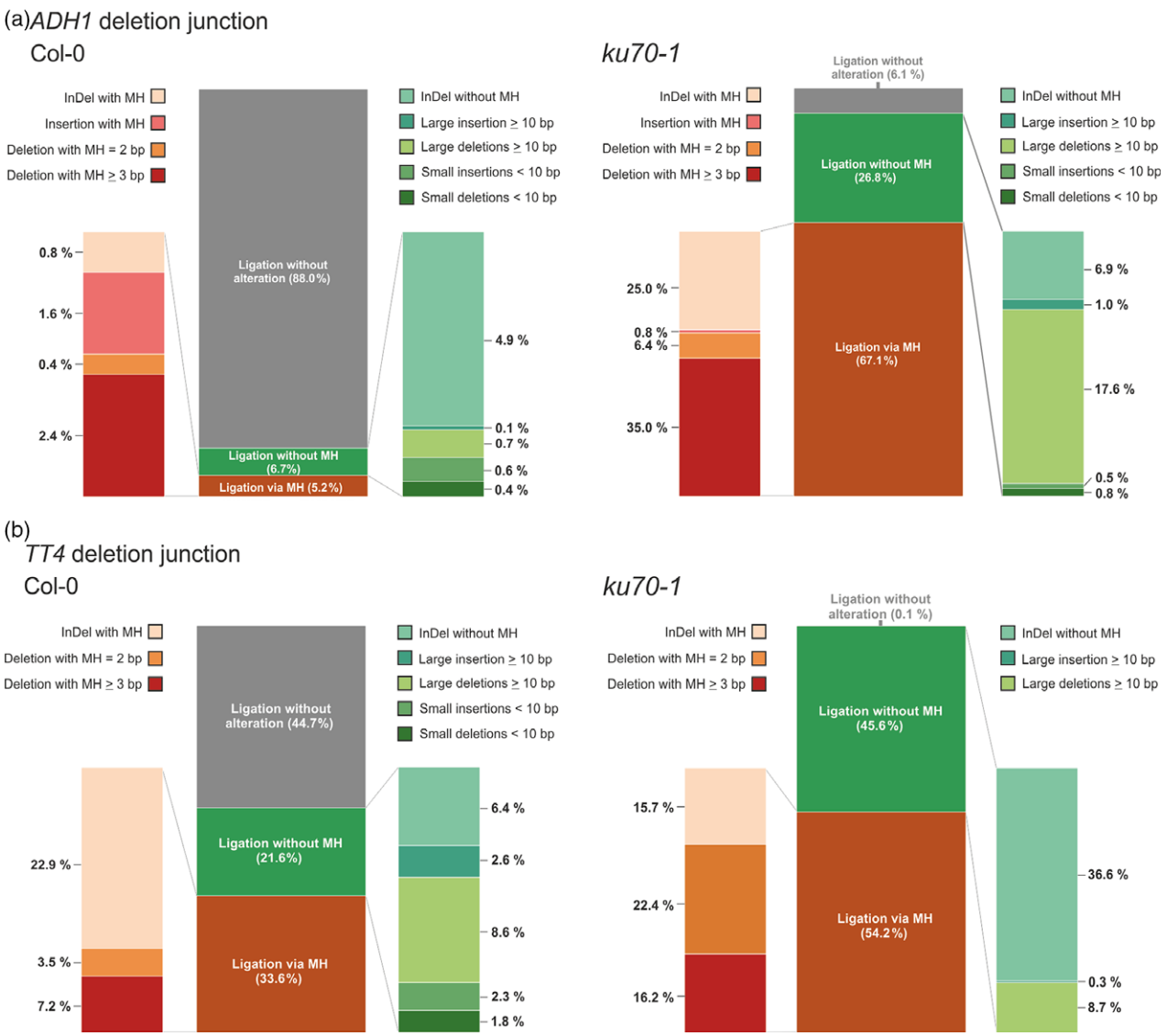


Figure 4. Detailed analysis of repair patterns of deletion junctions.
(a) Repair patterns of the deletion junction of the *ADH1* locus are pictured regarding the occurrence of micro-homologies (MHs) that were used for the formation of deletion junctions. We distinguished between precisely ligated junctions without alteration (gray), junctions repaired without the use of MHs (green) and repair events related to MH-mediated repair (orange). The different mutations were furthermore grouped within these classes due to their type (deletions, insertions, indels), size and MH size. In wild-type, most junctions showed precise ligation, the second-most frequent reads showed ligation site repair patterns without the use of MHs, and only a minority of junctions were repaired using MHs. In the *ku70-1* mutant, most of the junctions were repaired using MHs, and only a minority of the junctions were precisely ligated. About a quarter of all reads showed a repair pattern that is based on repair without the use of MHs. In contrast to wild-type, most of the junctions showed large deletions and indels formed via MHs.
(b) Repair patterns of the deletion junction of the *TT4* locus. In wild-type, about half of the junctions showed precise ligation, the second-most frequent reads showed ligation site repair patterns with the use of MHs, and about 20% of junctions were repaired without the use of MHs. In the *ku70-1* mutant, most of the junctions were repaired using MHs and only a minority of the junctions were precisely ligated. As for the *ADH1* deletion junctions, most of the junctions observed in *ku70-1* showed large deletions and indels.

background is not useful for practical applications. Therefore, we decided to set up a system applicable to wild-type plants, in which a very efficient transfer of these inversions to the next generation is achieved. We decided to use an egg-cell-specific promoter to express the Cas9/sgRNA system at a very early stage of plant development (Wang *et al.*, 2015). Recently, we were able to demonstrate that the efficiency of gene targeting by HR is significantly increased when using this promoter (Wolter *et al.*, 2018). Therefore, we exchanged the *Pc*Ubi4-2 promoter for the EC1.2en-EC1.1 promoter (EC1.2 enhancer fused to the EC1.1 promoter) combined with the rbcS-E9 terminator in the four above-mentioned constructs (*TT4*, *ADH1*, ECA3, *PSDG*). In a second set of experiments, we set out to induce inversions larger than 3 kb. For this purpose, we chose the *TT4* locus and used the same d protospacer for each construct, as with the *TT4* 3 kb approach, whilst the p protospacer was located at 8, 12 and 18 kb intervals. The seven constructs were transformed into *A. thaliana* wild-type plants via Agrobacterium-mediated transformation. The resultant selected seedlings were already analyzed in T1. We used up to 200 primary transformants and, as our goal was to identify those plants that had an inversion event during egg-cell transformation, the analysis was performed by simple genotyping via PCR (for primer combinations, see Table S5). Impressively, we identified at least one heritable inversion already in T1, in each tested approach and with frequencies up to 1.5% (Table 1).

Sanger sequencing verified the junctions of the T1 plants that were positively screened. The plants were then selfed, and the resulting progeny was tested for segregation of the induced inversion. Indeed, in all the cases tested, Mendelian segregation was found as expected for a hemizygous event, and homozygous inversion lines could be identified for each tested approach in T2. Therefore, viable fertile plants with induced inversions of up to almost 20 kb could be obtained with ease.

Table 1 Inversion frequencies of T1 plants using an egg-cell-specific promoter

Locus	Size	Plants tested in T1	Heritable inversions	T1 inversion efficiency (%)
ADH1	2.9 kb	199	2	1.01
TT4	3.4 kb	200	3	1.50
ECA3	2.8 kb	174	1	0.57
PSDG	3.1 kb	200	3	1.50
TT4 8 kb	8.3 kb	200	1	0.50
TT4 12 kb	12.9 kb	132	2	1.52
TT4 18 kb	18.2 kb	156	1	0.64

Results of PCR-based genotyping to identify inversion events in the T1 generation. For each approach up to 200 T1 plants were screened and there has been at least one heritable inversion event for each line.

Setup of a protocol for the efficient induction of heritable inversions

We considered how we could further improve the efficiency of obtaining heritable inversions. We speculated that stable expression of Cas9 in the egg cell, compared with transient transformation, might further enhance the chances to obtain heritable inversions. Following stable transformation, depending on the random integration site, as the expression of Cas9 might drastically vary between transgenic lines, it was important to develop a strategy to identify the most promising lines for screening.

We therefore grew 40 individual T1 plants out of the population that were not tested positive for the presence of an inversion for *ADH1*, *TT4*, *TT4* 8 kb and *TT4* 18 kb (Table 1). We selfed these plants and the T2 seeds were grown on GM media. Initially, in order to identify a large number of promising lines with as little effort as possible, a collective DNA extraction of 40 individual plants was done for each individual line. These bulk samples were screened for inversion events via PCR, and the lines that showed an easily detectable amount of PCR product of the expected size were examined further (Figure 5a). We screened 40 T2 lines each for both the p and d junctions of the inversion and found for *TT4* and *ADH1*, 14 and 10 T2 lines, respectively, that showed specific PCR products for both junctions. Additionally, we also tested the inversions spanning 8 and 18 kb based on the *TT4* approach and screened 20 T2 lines for each of these approaches. We were able to identify six T2 lines for the 8 kb inversion and four T2 lines for the 18 kb inversion. We then selected up to five positive lines for each kind of approach, and performed DNA extraction and PCR-based genotyping of 40 individual plants for each of these lines (Table 2).

For all lines, we found at least one plant harboring an inversion within their genome. The efficiencies ranged from 2.5 to 10.0%. On average, we were able to obtain plants carrying an inversion in the T2 in over 6% of cases for the 3 kb inversion, and in over 3% of cases for inversions larger than 8 kb.

For a more detailed characterization of the inversion we have chosen 12 T2 plants obtained from the *TT4* and 13 T2 from the *ADH1* locus, and sequenced both p and d junctions. For the inversions induced in the *ADH1* locus, all breaks were re-ligated perfectly to the break points for all 12 plants. In the case of the *TT4* locus, we detected 10 plants with perfectly ligated inversions; in two plants we found mutations at one junction and in one plant at both junctions (Figures 5d and S4c). Furthermore, we were interested in verifying that the inverted sequences did not acquire mutations or deletions during the inversion process. Primers were designed to allow specific amplification of the complete inverted sequence (Figures 5b and S4a). We were able to amplify all 25 inversions and verified the

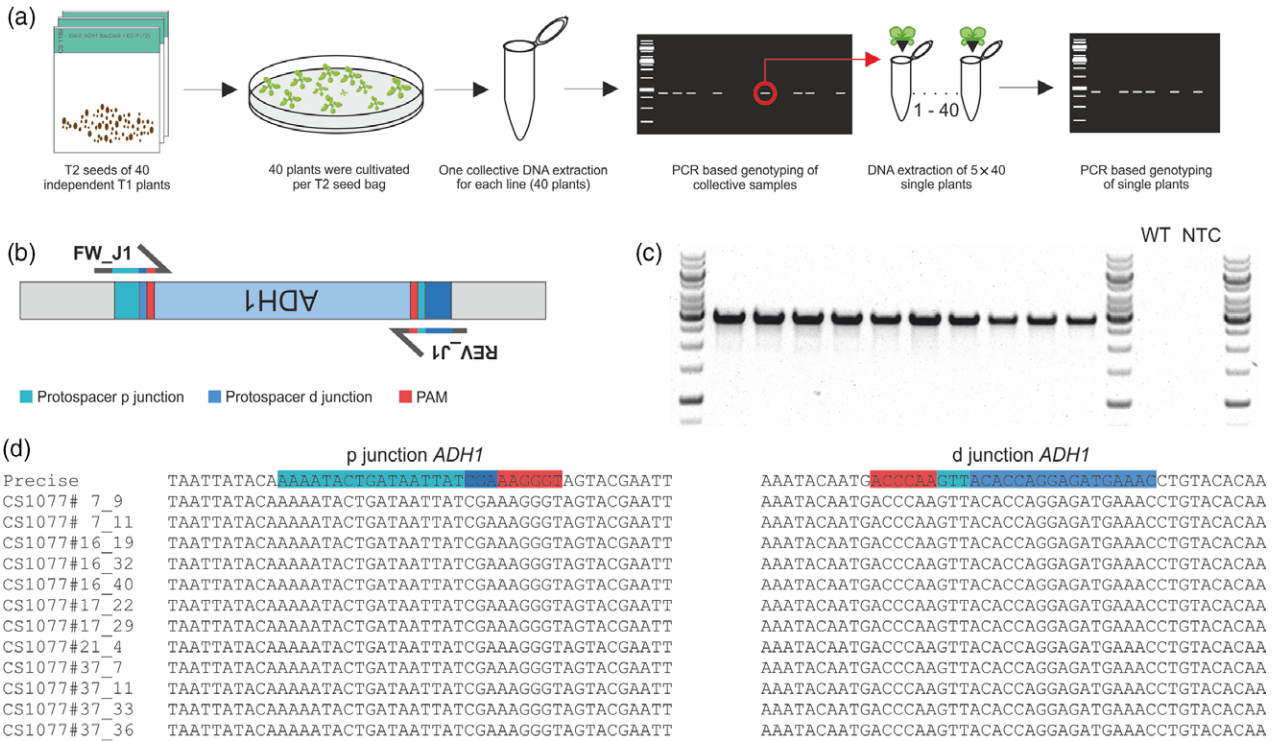


Figure 5. Analysis of the *Sa*Cas9-induced inversions obtained in the *ADH1* T2 generation.
(a) General setup of the screening strategy to quantify heritable inversion events in T2. Therefore, the T2 seeds of the 40 inconspicuous T1 lines were cultivated on germination medium (GM) agar plates, and a collective DNA extraction of one leaf per plant from each line was performed, resulting in 40 collective DNA samples. These collective DNA samples were then genotyped via polymerase chain reaction (PCR) to identify plant lines with several inversion events. Thereafter, those lines showing a distinct PCR band were analyzed in detail, and for this purpose DNA extraction was performed from each of the 40 plants.
(b) Schematic representation of the junction-specific primers used for the amplification of the entire inversion of the *ADH1* gene. Primers were designed to bridge the junctions to specifically enable inversion-specific amplification.
(c) Picture of gel electrophoresis of the PCR products after amplification of the complete inversion. All detectable bands are of the same size, indicating that no deletions occurred within the inverted sequences. As a control, we used the DNA of untransformed wild-type plants, as well as a no template control showing no bands, as expected.
(d) Sequencing results of both junctions of heritable inversions identified in T2. Each sample showed perfect ligation patterns at the p and d junctions.

expected size of the PCR products by gel electrophoresis (Figures 5c and S4b). Additionally, we performed Sanger sequencing of three randomly chosen samples for each approach. We detected no sequence alterations within the 3 kb inversions in all six sequenced samples. Thus, we could clearly show that Cas9-specific induction of inversions occurs almost exclusively error-free in Arabidopsis.

DISCUSSION

Till now, CRISPR/Cas technology has merely been used for gene-editing purposes in plants (Gao, 2018; Kumlehn *et al.*, 2018) and, in order to exploit its full potential for plant breeding, it is also important to set up technologies for various kinds of chromosomal engineering strategies (Puchta, 2017). By inducing two DSBs within a single chromosome, it is possible to obtain deletions or inversions. Although a number of papers on chromosomal deletion formation have been recently published (Ordon *et al.*, 2017; Durr *et al.*, 2018; Wu *et al.*, 2018), there has not been a study on efficient heritable inversion formation for plants till now. We demonstrated in this study that inversions up to 18 kb can be induced using the Cas9 ortholog from *Staphylococcus aureus* at different locations in the Arabidopsis genome with ease. Naturally occurring chromosomal rearrangements are observed repeatedly in many crop species, like wheat, rice, maize and tomato (Badaeva *et al.*, 2007; Fang *et al.*, 2012; Wang *et al.*, 2015; Rodríguez-Leal *et al.*, 2017). Inversions represent a major obstacle in plant breeding as recombination is inhibited in inverted regions. The current study indicates that it is possible to reverse natural inversions. This can enable the transfer of resistance markers from one cultivar to the other, as required, for example, to transfer the nematode-resistance gene Mi-1, which is associated with an inverted chromosomal segment in tomato (Seah *et al.*, 2004). Obviously, many natural inversions are much larger than the ones achieved in this study. Therefore, the next step will be to test our protocol on the induction of inversions in the Mbp range in Arabidopsis.

Table 2 Inversion frequencies of not previously modified T2 plants using an egg-cell-specific promoter

Locus	T2 plant line	Plants tested in T2	Heritable inversions	T2 inversion efficiency (%)
ADH1	CS 1107 # 7	40	2	5.0
	CS 1107 # 16	40	3	7.5
	CS 1107 # 17	40	2	5.0
	CS 1107 # 21	40	1	2.5
	CS 1107 # 37	40	4	10.0
TT4	CS 1099 # 1	40	3	7.5
	CS 1099 # 8	40	1	2.5
	CS 1099 # 20	40	4	10.0
	CS 1099 # 24	40	3	7.5
	CS 1099 # 32	40	2	5.0
TT4 8 kb	CS 1128 #2	40	2	5.0
	CS 1128 #4	40	1	2.5
	CS 1128 #5	40	2	5.0
	CS 1128 #6	40	2	5.0
	CS 1128 #7	40	1	2.5
TT4 18 kb	CS 1182 #2	40	1	2.5
	CS 1182 #13	40	2	5.0
	CS 1182 #15	40	1	2.5
	CS 1182 #18	40	1	2.5

Results of PCR-based genotyping to identify inversion events in the T2 generation using an egg-cell-specific promoter. Both *ADH1* and *TT4* inversions spanned about 3 kb. While for the *TT4* 8 and 18 kb inversions the same d protospacer was used as in the *TT4* approach, the p protospacer was moved to a distance of 8 and 18 kb. For each approach, up to five independent T2 lines with a total of 40 plants were tested. Heritable inversions were identified in each approach and frequencies of up to 10% were detected.

By determining the cutting efficiency using a T7 endonuclease assay (Figure S1), we were able to show that those protospacers with a high cutting efficiency also produced the largest amount of inversions. As it is a prerequisite that both DSBs are induced at the same time, inversion frequencies can be further enhanced by first carefully testing the protospacers used for break induction. Here, we not only demonstrate that the SaCas9 protein is suitable for the efficient induction of inversions in *A. thaliana*, but we also established a system to transfer these inversion events at high frequencies to the next generation.

Furthermore, we were able to define the mechanisms of inversion formation by analyzing the newly formed junctions on a molecular level. The majority of inversions obtained in wild-type showed precise ligation events without sequence changes at the newly formed junctions. Only a small number of mutated reads showed small deletions or insertions of a few bases. These repair patterns indicate that in the vast majority of cases inversions are formed by classical NHEJ (cNHEJ). The fact that junctions formed by simple ligation after inversion are not digestible by Cas9 anymore excludes that the enzyme is able to further modify the junctions by re-cutting. This is in contrast to the induction of a single DSB, where simple re-ligation of the junction would restore the cutting site and repeated re-digestion might lead to an active section of cutting-resistant mutated junctions. The observation that most Cas9-induced DSBs are actually repaired precisely in the context of cNHEJ, has also been recently reported in mammalian cells (Geisinger *et al.*, 2016; Guo *et al.*, 2018; Shou *et al.*, 2018), and our findings show that the same holds true for plants. To test if inversion formation was possible in the absence of cNHEJ, we induced inversions in a mutant lacking KU70, a key factor of cNHEJ. Astonishingly, we obtained enhanced frequencies of inversions, as opposed to reduced. The complete change in repair patterns, with a huge overrepresentation of MHs at the junctions, demonstrates that in the *ku70-1* mutant MH-mediated NHEJ is responsible for junction formation. Our results indicate that cNHEJ seems to have two different functions in DSB repair, not only to protect the broken ends to guarantee perfect ligation without sequence loss, but also to keep the two broken ends in close proximity to avoid chromosomal rearrangements and thus genome instabilities. The protein KU70 acts during cNHEJ together with KU80 as a heterodimer that binds to broken DNA ends and recruits further repair proteins (Tamura *et al.*, 2002; Lieber, 2010). Recent results in mammals indicate that KU recruited proteins like XLF and APLF are responsible for keeping both broken ends in close proximity to avoid genomic instability (Downs and Jackson, 2004; Graham *et al.*, 2018; Nemoz *et al.*, 2018). The presence of the KU heterodimer seems to prevent MH-mediated NHEJ factors like POLQ from processing DNA ends (Schimmel *et al.*, 2017). Thus, in the absence of KU, efficient MH-mediated NHEJ can take over. As the tethering of the 'right' ends is simultaneously lost, inversion frequencies are enhanced instead of being reduced. The fact that deletion formation under these conditions is also enhanced can be taken as strong support of our hypothesis. In addition to the loss of tethering, the coincidental presence of putative MHs might also influence the efficiency of inversion formation. However, this enhancement comes with a price: reduced accuracy. In the mutant background, we obtained large deletions and indels within many junctions. This makes the outcome of inversions in the *ku*-mutant much less predictable than in wild-type, due to the unforeseen further changes at the junctions. As *ku* mutants also show other genomic instabilities such as telomere dysfunction and DNA repair defects (Bundock *et al.*, 2002; Riha *et al.*, 2002), the use of this mutant is not advisable for practical applications in plant breeding. Therefore, we concentrated our further efforts on the generation of heritable events in wild-type plants. In order to achieve high heritable inversion efficiencies, we set the expression of the Cas9 ortholog from *S. aureus*, which has significantly higher cutting efficiencies than Cas9 of *Streptococcus pyogenes* in Arabidopsis (Steinert

et al., 2015), under the control of the egg-cell-specific promoter for tissue-specific expression, at an early stage of plant development. This promoter has already been used efficiently in Arabidopsis for the induction of heritable mutations by NHEJ, HR and also for the induction of deletions (Wang *et al.*, 2015; Durr *et al.*, 2018; Wolter *et al.*, 2018). Thus, we were able to generate heritable inversion events with frequencies in the percentage range already in T1, and these frequencies could be further increased in the T2 generation. These heritable inversions had in almost all cases junctions restored by simple re-ligation and without any mutations, similar to our results in somatic tissue. These findings indicate that the formation of inversions in egg cells does not differ significantly from inversions in somatic cells, and that comparable repair mechanisms are involved in both tissues. Moreover, by sequencing we checked a representative number of samples to exclude any rearrangements or the induction of mutations within the inverted fragments. As there was no case of any sequence change that could be detected, the applied methodology is error-free and the outcome is completely predictable. Predictable inversion formation in plant genomes could only previously be obtained by the use of site-specific recombinases (Medberry *et al.*, 1995). However, this methodology is strongly hampered, as it requires the insertion of specific recognition sequences in the plant genome at the respective sites prior to its application. Therefore, we are convinced that the CRISPR/Cas-based procedure described here represents a drastic improvement. Inversions can now be induced in a sequence-specific, highly accurate and effective way anywhere in plant genomes, making this approach attractive for various breeding applications.

EXPERIMENTAL PROCEDURES

Plant material and growth conditions

All *A. thaliana* lines used in this study were in the Columbia (Col-0) background and the T-DNA insertion line *ku70-1* (SALK_123114), from the SALK collection (Alonso *et al.*, 2003), was previously described (Jia *et al.*, 2012). Plants were cultivated in a growth chamber on agar plates containing germination medium (GM: 4.9 g L^{-1} Murashige & Skoog medium, 10 g L^{-1} saccharose, pH 5.7, 7.6 g L^{-1} plant agar) or in the greenhouse on substrate containing 1:1 mixture of Floraton 3 (Floragard, Oldenburg, Germany) and vermiculite (2–3 mm; Deutsche Vermiculite Dämmstoff, Sprockhövel, Germany), at 22°C with 16 h light and 8 h darkness.

T-DNA constructs

DNA constructs used in this study are based on the pDe-Sa-Cas9 and pEn-Sa-Chimera plasmids previously described (Steinert *et al.*, 2015, 2017), only the kanamycin resistance cassette was exchanged to a bar or gentamycin resistance cassette using the restriction enzymes *Pme*I and *Sbf*I. Both spacer sequences were cloned into individual pEn-Sa-Chimera vectors and integrated together into the pDe-Sa-Cas9. Spacer sequences used for cloning are listed in Table S1. While the first chimera was added using *Bsu36*I and *Mlu*I, the second chimera was transferred via a Gateway® LR-reaction, resulting in the final T-DNA constructs with a PcUbi4-2 promoter for constitutive expression and two U6-26-sgRNA cassettes for specific simultaneous induction of two DSBs. For egg-cell-specific expression, the EC1.1-promoter (composed of the EC1.1 promoter combined with the EC1.2 enhancer; Wang *et al.*, 2015) was amplified using AS84 and AS85, and the rbcS-E9 terminator was amplified using AS112 and AS113, from pHEE2E-TRI, and both fragments were inserted into the respective pDe-Sa-Cas9 vector via Gibson assembly® (New England Biolabs, NEB, https://www.neb.com/). The primers used in this study are listed in Table S2.

Plant transformation

Arabidopsis plants were transformed via floral dip using the Agrobacterium strain GV3101, as previously described (Clough and Bent, 1998).

Quantification via digital droplet polymerase chain reaction and amplicon deep sequencing

T1 seeds were sown on GM media containing cefotaxime and either phosphinotricin or gentamycin for 2 weeks. For each construct, 100 T1 plants were pooled and used for DNA extraction as previously described (Salomon and Puchta, 1998). Analysis via ddPCR was performed using site-specific primers (Table S2) and dual-labelled probes (Table S3). Both junctions and the control were measured using probe and EvaGreen®-based assays (Table S4). The ddPCR was performed using the QX200TM AutoDGTM Droplet DigitalTM PCR system, reagents, plates/cartridges from BioRad, and the following analysis was performed using the QuantasoftTM Analysis Pro software from BioRad. The extracted DNA of the 100 T1 plants was also used for amplicon deep sequencing. Therefore, primers were designed with a 6-bp tag sequence (Table S2) to amplify amplicons with a length of 370–400 bp using the Q5® High-Fidelity DNA Polymerase (New England Biolabs, NEB, https://www.neb.com/). These amplicons were purified using the peqGOLD Cycle-Pure Kit (Peqlab), and afterwards sequenced using the Illumina HiSeq platform at GATC Biotech. After tag sorting, the data were analyzed using the CRISPR RGEN Tool (Park *et al.*, 2017), R and Excel.

T7 endonuclease assay

To determine genome-targeting efficiency, we used the T7 Endonuclease I (New England Biolabs, NEB, https://www.neb.com/), which can recognize and cleave non-perfectly matched DNA. Therefore, we designed primers spanning the respective protospacer sequence with a distance of about 1 kb. We amplified the regions of both protospacers using the DNA of T1 plant pools (the same pools used for ddPCR and amplicon deep sequencing) and an untransformed wild-type control with the Q5® High-Fidelity DNA Polymerase (New England Biolabs, NEB, https://www.neb.com/). After purification via the peqGOLD Cycle-Pure Kit (Peqlab, VWR International GmbH, Darmstadt, Germany), PCR products were annealed and subsequently digested with T7 Endonuclease I, according to the NEB protocol. (https://international.neb.com/protocols/2014/08/11/determining-genome-targeting-efficiency-using-t7-endonuclease-i). Fragments were then separated by agarose gel electrophoresis, visualized with UV light and analyzed.

DECLARATIONS

Availability of data and material

The datasets generated and/or analyzed during the current study are available in the NCBI Sequence Read Archive (SRA) repository, BioProject ID PRJNA516353 (http://www.ncbi.nlm.nih.gov/bioproject/516353).

FUNDING

This work was funded by the European Research Council ERC [https://erc.europa.eu; grant numbers ERC-2016-AdG_741306 CRISBREED].

ACKNOWLEDGEMENTS

The authors wish to thank Stefanie Wunderlich, Carina Jülch and Waltraud Wehrle for excellent technical assistance, Amy Whitbread, Annika Dorn and Felix Wolter for critically reading the manuscript, and Tobias Zundel for developing an R-based bioinformatic solution for detailed analysis of deep sequencing results.

CONFLICT OF INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

CS, MP and HP designed research; CS performed research; CS analyzed data; and CS and HP wrote the paper.

SUPPORTING INFORMATION

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

Figure S1. Results of the T7 endonuclease assay to determine cutting efficiencies.

Figure S2. Length of MHs used for repair at the *TT4* locus on both inversion junctions in wild-type.

Figure S3. Detailed analysis of repair patterns of both *TT4* inversion junctions.

Figure S4. Analysis of the SaCas9-induced inversions obtained in the *TT4* T2 generation.

Table S1. Spacer oligonucleotide sequences with overhangs.

Table S2. Primer sequences.

Table S3. LNA probe sequences.

Table S4. Primer combinations used for ddPCR.

Table S5. Primer combination for PCR-based genotyping.

REFERENCES

Alonso, J.M., Stepanova, A.N., Leisse, T.J. *et al.* (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science* **301**, 653–657.

Badaeva, E.D., Dedkova, O.S., Gay, G., Pukhalskyi, V.A., Zelenin, A.V., Bernard, S. and Bernard, M. (2007) Chromosomal rearrangements in wheat: their types and distribution. *Genome* **50**, 907–926.

Blanc, G., Barakat, A., Guyot, R., Cooke, R. and Delseny, M. (2000) Extensive duplication and reshuffling in the Arabidopsis genome. *Plant Cell* **12**, 1093–1101.

Bundock, P., van Attikum, H. and Hooykaas, P. (2002) Increased telomere length and hypersensitivity to DNA damaging agents in an Arabidopsis KU70 mutant. *Nucleic Acids Res.* **30**, 3395–3400.

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.

Cheong, T.-C., Blasco, R.B. and Chiarle, R. (2018) The CRISPR/Cas9 system as a tool to engineer chromosomal translocation in vivo. *Adv. Exp. Med. Biol.* **1044**, 39–48.

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. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823.

Downs, J.A. and Jackson, S.P. (2004) A means to a DNA end: the many roles of Ku. *Nat. Rev. Mol. Cell Biol.* **5**, 367–378.

Durr, J., Papareddy, R., Nakajima, K. and Gutierrez-Marcos, J. (2018) Highly efficient heritable targeted deletions of gene clusters and non-coding regulatory regions in Arabidopsis using CRISPR/Cas9. *Sci. Rep.* **8**, 4443.

Fang, Z., Pyhäjärvi, T., Weber, A.L., Dawe, R.K., Glaubitz, J.C., González Jde, J.S., Ross-Ibarra, C., Doebley, J., Morrell, P.L. and Ross-Ibarra, J. (2012) Megabase-scale inversion polymorphism in the wild ancestor of maize. *Genetics* **191**, 883–894.

Fransz, P., Linc, G., Lee, C.-R. *et al.* (2016) Molecular, genetic and evolutionary analysis of a paracentric inversion in Arabidopsis thaliana. *Plant J.* **88**, 159–178.

Gao, C. (2018) The future of CRISPR technologies in agriculture. *Nat. Rev. Mol. Cell Biol.* **19**, 275–276.

Geisinger, J.M., Turan, S., Hernandez, S., Spector, L.P. and Calos, M.P. (2016) In vivo blunt-end cloning through CRISPR/Cas9-facilitated non-homologous end-joining. *Nucleic Acids Res.* **44**, e76.

Graham, T.G.W., Carney, S.M., Walter, J.C. and Loparo, J.J. (2018) A single XLF dimer bridges DNA ends during nonhomologous end joining. *Nat. Struct. Mol. Biol.* **25**, 877–884.

Guo, T., Feng, Y.-L., Xiao, J.-J. *et al.* (2018) Harnessing accurate non-homologous end joining for efficient precise deletion in CRISPR/Cas9-mediated genome editing. *Genome Biol.* **19**, 170.

Jia, Q., Bundock, P., Hooykaas, P.J.J. and de Pater, S. (2012) Agrobacterium tumefaciens T-DNA integration and gene targeting in Arabidopsis thaliana non-homologous end-joining mutants. *J. Bot.* **2012**, 13.

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.

Kirkpatrick, M. and Barton, N. (2006) Chromosome inversions, local adaptation and speciation. *Genetics*, **173**, 419–434.

Kumlehn, J., Pietralla, J., Hensel, G., Pacher, M. and Puchta, H. (2018) The CRISPR/Cas revolution continues: from efficient gene editing for crop breeding to plant synthetic biology. *J. Integr. Plant Biol.* **60**, 1127–1153.

Li, W., Challa, G.S., Zhu, H. and Wei, W. (2016) Recurrence of chromosome rearrangements and reuse of DNA breakpoints in the evolution of the Triticeae genomes. *G3 (Bethesda)* **6**, 3837–3847.

Lieber, M.R. (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79**, 181–211.

Lin, T., Zhu, G., Zhang, J. *et al.* (2014) Genomic analyses provide insights into the history of tomato breeding. *Nat. Genet.* **46**, 1220–1226.

Lowry, D.B. and Willis, J.H. (2010) A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.* **8**, e1000500.

Madan, K. (1995) Paracentric inversions: a review. *Hum. Genet.* **96**, 503–515.

Medberry, S.L., Dale, E., Qin, M. and Ow, D.W. (1995) Intra-chromosomal rearrangements generated by Cre-lox site-specific recombination. *Nucleic Acids Res.* **23**, 485–490.

Navarro, A. and Barton, N.H. (2003) Chromosomal speciation and molecular divergence-accelerated evolution in rearranged chromosomes. *Science* **300**, 321–324.

Nemoz, C., Ropars, V., Frit, P. *et al.* (2018) XLF and APLF bind Ku80 at two remote sites to ensure DNA repair by non-homologous end joining. *Nat. Struct. Mol. Biol.* **25**, 971–980.

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.

Pacher, M., Schmidt-Puchta, W. and Puchta, H. (2007) Two unlinked double-strand breaks can induce reciprocal exchanges in plant genomes via homologous recombination and nonhomologous end joining. *Genetics* **175**, 21–29.

Park, J., Lim, K., Kim, J.-S. and Bae, S. (2017) Cas-analyzer: an online tool for assessing genome editing results using NGS data. *Bioinformatics* **33**, 286–288.

Puchta, H. (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *J. Exp. Bot.* **56**, 1–14.

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. (2014) Synthetic nucleases for genome engineering in plants: prospects for a bright future. *Plant J.* **78**, 727–741.

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.

Qi, Y., Li, X., Zhang, Y., Starker, C.G., Baltes, N.J., Zhang, F., Sander, J.D., Reyon, D., Joung, J.K. and Voytas, D.F. (2013a) Targeted deletion and inversion of tandemly arrayed genes in Arabidopsis thaliana using zinc finger nucleases. *G3 (Bethesda)* **3**, 1707–1715.

Qi, Y., Zhang, Y., Zhang, F., Baller, J.A., Cleland, S.C., Ryu, Y., Starker, C.G. and Voytas, D.F. (2013b) Increasing frequencies of site-specific mutagenesis and gene targeting in Arabidopsis by manipulating DNA repair pathways. *Genome Res.* **23**, 547–554.

Riha, K., Watson, J.M., Parkey, J. and Shippen, D.E. (2002) Telomere length deregulation and enhanced sensitivity to genotoxic stress in Arabidopsis mutants deficient in Ku70. *EMBO J.* **21**, 2819–2826.

Rodríguez-Leal, D., Lemmon, Z.H., Man, J., Bartlett, M.E. and Lippman, Z.B. (2017) Engineering quantitative trait variation for crop improvement by genome editing. *Cell* **171**, 470–480.e8.

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.

Schimmel, J., Kool, H., van Schendel, R. and Tijsterman, M. (2017) Mutational signatures of non-homologous and polymerase theta-mediated end-joining in embryonic stem cells. *EMBO J.* **36**, 3634–3649.

Schubert, I. (2018) What is behind "centromere repositioning"? *Chromosoma* **127**, 229–234.

Schubert, I. and Vu, G.T.H. (2016) Genome stability and evolution: attempting a holistic view. *Trends Plant Sci.* **21**, 749–757.

Seah, S., Yaghoobi, J., Rossi, M., Gleason, C.A. and Williamson, V.M. (2004) The nematode-resistance gene, Mi-1, is associated with an inverted chromosomal segment in susceptible compared to resistant tomato. *Theor. Appl. Genet.* **108**, 1635–1642.

Shen, H., Strunks, G.D., Klemann, B.J.P.M., Hooykaas, P.J.J. and de Pater, S. (2017) CRISPR/Cas9-induced double-strand break repair in Arabidopsis nonhomologous end-joining mutants. *G3 (Bethesda)* **7**, 193–202.

Shou, J., Li, J., Liu, Y. and Wu, Q. (2018) Precise and predictable CRISPR chromosomal rearrangements reveal principles of Cas9-mediated nucleotide insertion. *Mol. Cell* **71**, 498–509.e4.

Siebert, R. and Puchta, H. (2002) Efficient repair of genomic double-strand breaks by homologous recombination between directly repeated sequences in the plant genome. *Plant Cell* **14**, 1121–1131.

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., Schmidt, C. and Puchta, H. (2017) Use of the Cas9 orthologs from Streptococcus thermophilus and Staphylococcus aureus for non-homologous end-joining mediated site-specific mutagenesis in Arabidopsis thaliana. *Methods Mol. Biol.* **1669**, 365–376.

Szinay, D., Wijnker, E., van den Berg, R., Visser, R.G.F., de Jong, H. and Bai, Y. (2012) Chromosome evolution in Solanum traced by cross-species BAC-FISH. *New Phytol.* **195**, 688–698.

Tamura, K., Adachi, Y., Chiba, K., Oguchi, K. and Takahashi, H. (2002) Identification of Ku70 and Ku80 homologues in Arabidopsis thaliana: evidence for a role in the repair of DNA double-strand breaks. *Plant J.* **29**, 771–781.

Udall, J.A., Quijada, P.A. and Osborn, T.C. (2005) Detection of chromosomal rearrangements derived from homologous recombination in four mapping populations of Brassica napus L. *Genetics* **169**, 967–979.

Voytas, D.F. (2013) Plant genome engineering with sequence-specific nucleases. *Annu. Rev. Plant Biol.* **64**, 327–350.

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.

Wolter, F., Klemm, J. and Puchta, H. (2018) Efficient in planta gene targeting in Arabidopsis using egg cell-specific expression of the Cas9 nuclease of Staphylococcus aureus. *Plant J.* **94**, 735–746.

Wu, R., Lucke, M., Jang, Y.-T., Zhu, W., Symeonidi, E., Wang, C., Fitz, J., Xi, W., Schwab, R. and Weigel, D. (2018) An efficient CRISPR vector toolbox for engineering large deletions in Arabidopsis thaliana. *Plant Methods* **14**, 65.

Zapata, L., Ding, J., Willing, E.-M. *et al.* (2016) Chromosome-level assembly of Arabidopsis thaliana Ler reveals the extent of translocation and inversion polymorphisms. *Proc. Natl Acad. Sci. USA* **113**, E4052–E4060.

Zhang, C., Liu, C., Weng, J., Cheng, B., Liu, F., Li, X., Xie, C. (2017) Creation of targeted inversion mutations in plants using an RNA-guided endonuclease. *Crop J.* **5**, 83–88.

