Role of non-homologous end-joining in T-DNA integration in Arabidopsis thaliana

Shen, H.

Chapter 5

Sequence-specific nuclease-induced double strand break repair in Arabidopsis non-homologous end-joining mutants

Hexi Shen*, Gary D Strunks*, Bart J. P. M. Klemann, Paul J. J. Hooykaas, Sylvia de Pater

Department of Molecular and Developmental Genetics, Institute of Biology, Leiden University, Leiden, 2333 BE, The Netherlands

Part of this work is published as
H. Shen*, G. D. Strunks*, B. J. P. M. Klemann, P. J. J. Hooykaas and S. de Pater. (2017) CRISPR/Cas9-induced double-strand break repair in Arabidopsis nonhomologous end-joining mutants. *G3 Genes Genomes genetics* doi: 10.1534/g3.116.035204.

Abstract

Double-strand breaks (DSBs) are one of the most harmful DNA lesions. Cells utilize two main pathways for DSB repair: homologous recombination (HR) and non-homologous end joining (NHEJ). NHEJ can be subdivided into the KU-dependent classical NHEJ (c-NHEJ) and the more error-prone KU-independent backup-NHEJ (b-NHEJ) pathways, involving the poly (ADP-ribose) polymerases (PARPs). However, in absence of these factors, cells still seem able to adequately maintain genome integrity, suggesting the presence of other b-NHEJ repair factors or pathways independent from KU and PARPs. The outcome of DSB repair by NHEJ pathways can be investigated by using artificial sequence-specific nucleases such as TALENs and CRISPR/Cas9 to induce DSBs at a target of interest. Here, we used TALEN and CRISPR/Cas9 for DSB induction at the *Arabidopsis cruciferin 3* (*CRU3*) and *protoporphyrinogen oxidase* (*PPO*) genes. DSB repair outcomes via NHEJ were analysed using footprint analysis in wild-type plants and plants deficient in the c-NHEJ pathway (*ku80*), the b-NHEJ pathway (*parp1parp2*) or both (*ku80parp1parp2*). We found that larger deletions of more than 20 bp predominated after DSB repair in *ku80* and *ku80parp1parp2* mutants, corroborating with a role of KU in preventing DSB end resection. Deletion lengths did not significantly differ between *ku80* and *ku80parp1parp2* mutants, suggesting that a KU and PARP-independent b-NHEJ mechanism becomes active in these mutants. Furthermore, microhomologies and templated insertions were observed at the repair junctions in the wild type and all mutants. Since these characteristics are hallmarks of Polymerase θ-mediated DSB repair, we suggest a possible role for this recently discovered polymerase in DSB repair in plants.

Introduction

Double-strand breaks (DSBs) are one of the most lethal forms of DNA damage. DSBs can occur during normal cellular metabolism or can be induced by external factors, and highly threaten genomic integrity and cell survival (Deriano and Roth 2013). To repair DSBs, cells have two main pathways: Homologous Recombination (HR) and Non-Homologous End-Joining (NHEJ). Both of them function together to maintain genome integrity. NHEJ is the predominant pathway in higher eukaryotes and repair may lead to mutations at break sites, such as deletions, insertions and substitutions. At least two NHEJ pathways have been identified: the classic NHEJ pathway (c-NHEJ) and the backup-NHEJ pathway (b-NHEJ) also called alternative-NHEJ (a-NHEJ) or microhomology-mediated end-joining (MMEJ). The c-NHEJ is initiated by the recognition and binding of the KU heterodimer, consisting of KU70 and KU80 subunits, to DSBs (Walker *et al.* 2001). Once bound to a DSB, the KU heterodimer serves as a scaffold to recruit other c-NHEJ factors to the broken ends and promotes end-joining. Because KU is the key component of the c-NHEJ pathway, this pathway is also called KU-dependent NHEJ. In the absence of KU, other factors gain entry to the DSB site for repair by backup pathways. Although the b-NHEJ pathway was defined by involving multiple components, including poly (ADP-ribose) polymerase 1 (PARP1) the precise mechanism is still not clear (Wang *et al.* 2006). Furthermore, recently PARP1 was shown to be involved in repair of DSBs also in the presence of KU (Luijsterburg *et al.* 2016).

Nowadays, DSBs can be induced artificially at specific sites in the genome by sequence-specific artificial nucleases, which can be used to study DSB repair. These induced DSBs will be mainly repaired via NHEJ, which may lead to targeted mutagenesis. When repair restores the target site for the nuclease, the sequence will be cut again in the continuous presence of the nuclease. This cycle of cutting and repair will continue until incorrect repair destroys the target site. When a homologous sequence such as a sister chromatid, is present, DSB repair may also occur via HR, but this will inevitably also lead to restoration of the target site. A repair template without the target site may be provided by transformation or pre-inserted in the genome, and, when used for repair, lead to gene targeting (GT) (Voytas 2013; Puchta and Fauser 2013a). The current genome editing tool kit comprises four classes of engineered nucleases: modified meganucleases, zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and the CRISPR/Cas9 (clustered regularly interspaced short palindromic repeat/CRISPR-associated 9) system (Voytas 2013; Puchta and Fauser 2013a), of which the TALENs and CRISPR/Cas9 system are the most easy and straightforward to use.

TALENs consist of a DNA binding domain, derived from proteins produced by plant pathogens of the genus *Xanthomonas*, fused to the *Fok*I nuclease domain, and they can cleave DNA as a dimer (Christian *et al.* 2010). The DNA binding domain of TALENs consists of an array of 13 – 28 repeats. Each repeat consists of 34 highly conserved amino acids, of which only the amino acid residues at position 13 and 14, also called repeat variable diresidues (RVDs), vary and can specifically bind to one of each of the four DNA bases (Voytas 2013). In this way, an array of TAL effector repeats has a one to one correspondence with the DNA sequence it binds. A TALEN pair can recognize 26 – 56 bp, a sequence length which can be considered unique in higher eukaryotic genomes. The efficiency of DSB induction by

TALENs is also influenced by the spacer length between the binding sites of the two TAL arrays (Christian *et al.* 2010).

CRISPR/Cas9 is the most recent addition to the genome editing tool box (Jinek *et al.* 2012). It is derived from an adaptive immune system present in bacteria and archaea, where it serves in degrading invading foreign plasmid or viral DNA. The type II CRISPR genomic locus encodes the Cas9 ('CRISPR-associated 9') endonuclease, which can form a complex with two short RNA molecules: CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA), which guide the Cas9 protein to a DNA sequence of interest where it can induce a DSB through cleavage of the two DNA strands with its two nuclease domains (RuvC-like domain I and HNH motif). It was shown that the crRNA and tracrRNA can be fused into a chimeric single-guide RNA (sgRNA) comprising the functions of both precursor RNAs (Jinek *et al.* 2012). A sgRNA can be assembled to target any DNA sequence, with the prerequisite that a protospacer adjacent motif (PAM) sequence of NGG flanking the 3'end of the sgRNA target sequence is present, which interacts with the Cas9 PAM interacting domain (PI domain) (Nishimasu *et al.* 2014; Jinek *et al.* 2014). The direct RNA-DNA recognition of the CRISPR/Cas9 system has the advantage that only the sequence of the sgRNA needs to be changed if new loci have to be targeted, instead of the more laborious assembly of new TAL effector arrays when using TALENs. Furthermore, CRISPR/Cas9 conveniently allows DSB induction at multiple targets by simultaneously expressing multiple sgRNAs in combination with the Cas9 protein (Cong *et al.* 2013).

Previous studies already demonstrated the feasibility of DSB-mediated targeted mutagenesis at artificial and endogenous loci in plants using ZFNs, TALENs and the CRISPR/Cas9 system (Puchta and Fauser 2013b). Here, we used TALENs and CRISPR/Cas9 for DSB-mediated targeted mutagenesis at the *Arabidopsis cruciferin 3* (*CRU3*) and *protoporphyrinogen oxidase* (*PPO*) genes. CRISPR/Cas9 nucleases were expressed in mutants in the c-NHEJ or b-NHEJ DNA repair pathways and a combination of both. Footprint analysis in whole seedlings in the wild type and each of the three mutant genotype backgrounds (*ku80, parp1parp2 and ku80parp1parp2* mutants) demonstrated that key factors of NHEJ can affect the outcomes of targeted mutagenesis.

Materials and Methods

Plant material

The *ku80* (SALK_016627), *parp1* (GABI-Kat Line 692A05) and *parp2* (SALK_140400) T-DNA insertion lines (ecotype Col-0), the *parp1parp2* double mutant and *ku80parp1parp2* triple mutant were described previously (Jia *et al.* 2013). More information about these lines can be found at http://signal.salk.edu/cgi-bin/tdnaexpress (Alonso *et al.* 2003).

TALEN and CRISPR/Cas9 vector construction and plant transformation

TALEN was designed and assembled with the Golden Gate Kit (AddGene) as described (Cermak *et al.* 2011). Arrays of repeats with the corresponding RVDs (**Table 1**) for the DNA binding sequences of the *CRU3* target were assembled together in vector pZHY500 (TALEN-CRU-1-left) or pZHY501 (TALEN-CRU-1-right). Individual TALE repeats were cloned into vector pZHY013 using *Xba*I and *BamH*I (TALEN-CRU-1-left) or *Nhe*I and *Bgl*II (TALEN-CRU-1-right). Subsequently, the DNA fragment containing the TALEN pair was cloned in the binary vector pMDC32 (Curtis and Grossniklaus 2003), a 35S T-DNA expression vector, via a Gateway LR reaction to create TALEN-CRU-1 (pSDM3906).

For the CRISPR/Cas9 constructs, oligo's SP509 and SP510 (*CRU3* target) and SP512 and SP513 (*PPO* target) (**Table 2**) were annealed and cloned in *Bbs*I digested pEn-Chimera (Fauser *et al.* 2014). Subsequently, gene coding sgRNAs were cloned in pDE-pUbi-Cas9 (Fauser *et al.* 2014) by a Gateway LR reaction, resulting in Cas9-PPO (pSDM3905) and Cas9-CRU (pSDM3903), respectively.

Plant binary vectors were introduced into *Agrobacterium tumefaciens* strain AGL1 by electroporation. *Arabidopsis thaliana* plants of the Col-0 ecotype (wild type, *ku80*, *parp1parp2*, *ku80parp1parp2)* were transformed with T-DNAs containing nuclease expression cassettes, using the floral dip method (Clough and Bent 1998). T1 seeds were grown on MA solid medium without sucrose, supplemented with timentin (100 µg/mL), nystatin (100 µg/mL) and antibiotics for T-DNA selection: 15 µg/mL hygromycin for TALEN; 15 µg/mL phosphinothricin for CRSIPR/Cas9.

DNA isolation and footprint analysis

T2 seeds were germinated on ½ MS with T-DNA selection, the seedlings were disrupted to a powder under liquid N_2 in a Tissuelyser (Retch, Haan, Germany). Genomic DNA was extracted by the CTAB method (De Pater *et al.* 2009). For predigestion, one µg of genomic DNA was digested with *Dde*I (for TALEN-CRU-1 analysis), *Pst*I (Cas9-CRU analysis) or *Fau*I (for Cas9-PPO analysis) overnight and precipitated. Undigested or predigested DNA was used for PCR with Phusion polymerase (Thermo Scientific) to amplify the nucleases target sites, followed by digestion of the PCR products with *Dde*I, *Pst*I or *Fau*I and separated in agarose gels. PCR primers are shown in **Table 2**. Primers SP491 and SP492 were used for amplification of the TALEN-CRU-1 target region, primers SP245 and SP248 were used for the Cas9-CRU target region and primers SP392 and SP538 were used for the Cas9-PPO target region. The resistant fragments were isolated from gel and cloned into pJet1.2 (Thermo Scientific) and sequenced by Macrogen Europe (Amsterdam, The Netherlands). Identical sequences in the same line were considered as one mutagenesis event since they might have

been preselected from PCR amplification. Two-tailed Mann-Whitney tests were performed for statistical analysis of deletion- and insertion lengths .

Measurements of the mutation rate

To estimate the rate of TALEN- and Cas9-induced mutations the target sites were amplified using undigested genomic DNA. PCR products were digested with the appropriate restriction enzymes and analysed on agarose gels. The intensity of bands was quantified using ImageJ software. The mutation rate was calculated by dividing the intensity of the digest-resistant band by the total intensity of all bands in a given lane (Nekrasov *et al.* 2013).

High resolution melting

High resolution melting (HRM) analyses were performed on PCR clones from T2 seedlings of wild-type lines Cas9-CRU #2 and Cas9-PPO #7 using Precision Melt Supermix (Bio-Rad), containing EvaGreen saturated dye, and the Bio-Rad C1000 Touch thermal cycler (Bio-Rad). Melt curves were analyzed using the Bio-Rad Precision Melt Analysis software. For the *CRU* target primers SP492 and SP563 were used *and for the PPO* target primers SP560 and SP561 (**Table 2**). Samples that showed melt curve differences were sequenced by Macrogen Europe (Amsterdam, The Netherlands).

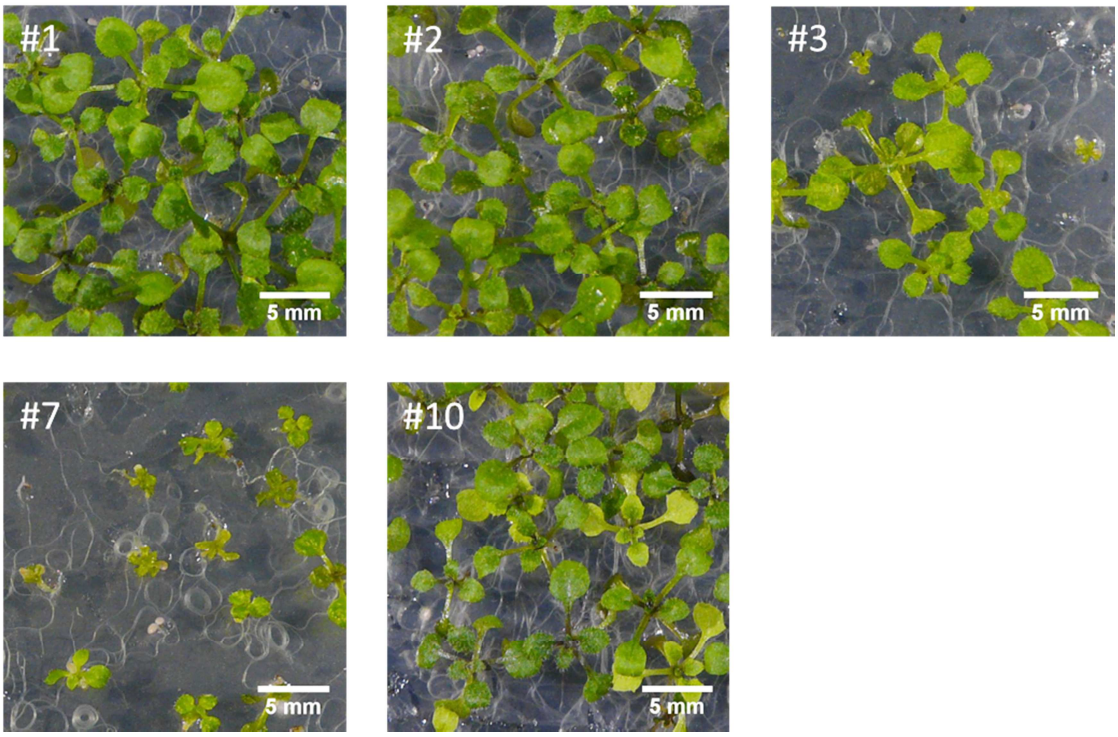


Figure 1. Phenotypes of 16 day-old T2 seedlings of 5 independent Cas9-*PPO* transformants. A stunted growth phenotype is observed in some seedlings of line #3 and #7. The other lines have a phenotype similar to wild-type.

Results

DSB-mediated mutagenesis by TALEN and CRISPR/Cas9 at the CRU3 and PPO loci

In order to investigate repair of induced DSBs, several sequence-specific nucleases were designed and expressed in *Arabidopsis*. Wild-type plants were transformed with TALEN or CRISPR/Cas9 expression constructs via the *Agrobacterium*-mediated floral dip method (Clough and Bent 1998) and T2 transformants were used for further analysis. Nuclease target sites in the *CRU3* and *PPO* genes were selected. The *CRU3* gene encodes a seed storage protein. The *PPO* gene encodes an essential enzyme that is involved in the final step of chlorophyll biosynthesis, and mutagenesis of the *PPO* gene is therefore toxic to plants. Plants expressing nucleases targeted at *CRU3* showed a phenotype similar to wild-type, but T2 seedlings of some plant lines expressing Cas9-PPO showed a stunted growth phenotype indicative of homozygous inactivation of the essential *PPO* gene in many cells (**Figure 1**). To detect mutagenesis caused by nuclease activity and subsequent erroneous NHEJ-mediated DSB repair at the molecular level, genomic DNA from T2 seedlings was analyzed for the presence of NHEJ-induced indels. In order to discriminate DNA molecules with a mutation, PCR products from the region containing the target site were digested with restriction enzymes having a recognition site overlapping or near the DSB site (*Dde*I for TALEN-CRU-1, *Pst*I for Cas9-CRU, and *Fau*I for Cas9-PPO) (**Figure 2**). Loss of the restriction site as a consequence of erroneous repair resulted in restriction digest-resistant PCR products. After gel electrophoresis, the relative band intensities were measured to estimate the mutation frequency in the target sites (**Figure 3**). Digestion of the PCR products from untransformed wild-type plants only left up to 3% of the material undigested, probably due to incomplete digestion. However, a distinguishably higher fraction of the PCR products from plant lines transformed with either TALEN or CRISPR/Cas9 nucleases were resistant for enzyme digestion. An average mutation rate of 5.5% of was found in TALEN-CRU-1 lines, 6.2% in Cas9-CRU lines and 11.2% in Cas9-PPO lines (**Figure 3**).

To get a better insight into the mutations induced by the TALEN and CRISPR/Cas9 nucleases, the restriction digest-resistant PCR products of the *CRU3* and *PPO* targets were cloned and sequenced. Pre-digested genomic DNA was used for PCR to enrich for mutated sequences. Sequencing revealed mainly deletions and some insertions and substitutions in both TALEN and CRISPR/Cas9 lines (**Figure 3D**, **Figure 7**). TALEN-CRU-1 footprints mainly consisted of small deletions ranging from 1 bp to 15 bp. The main reason we detected only small deletions in the TALEN lines may have been the small size of the amplified PCR fragments, due to the presence of additional *Dde*I sites just outside the amplified region. CRISPR/Cas9 lines showed more frequent and generally larger deletions compared to the TALEN lines (**Figure 3D**). Short homologous sequences on either the left or right side flanking the deletion were often also present, suggesting MMEJ may have been involved in DSB repair.

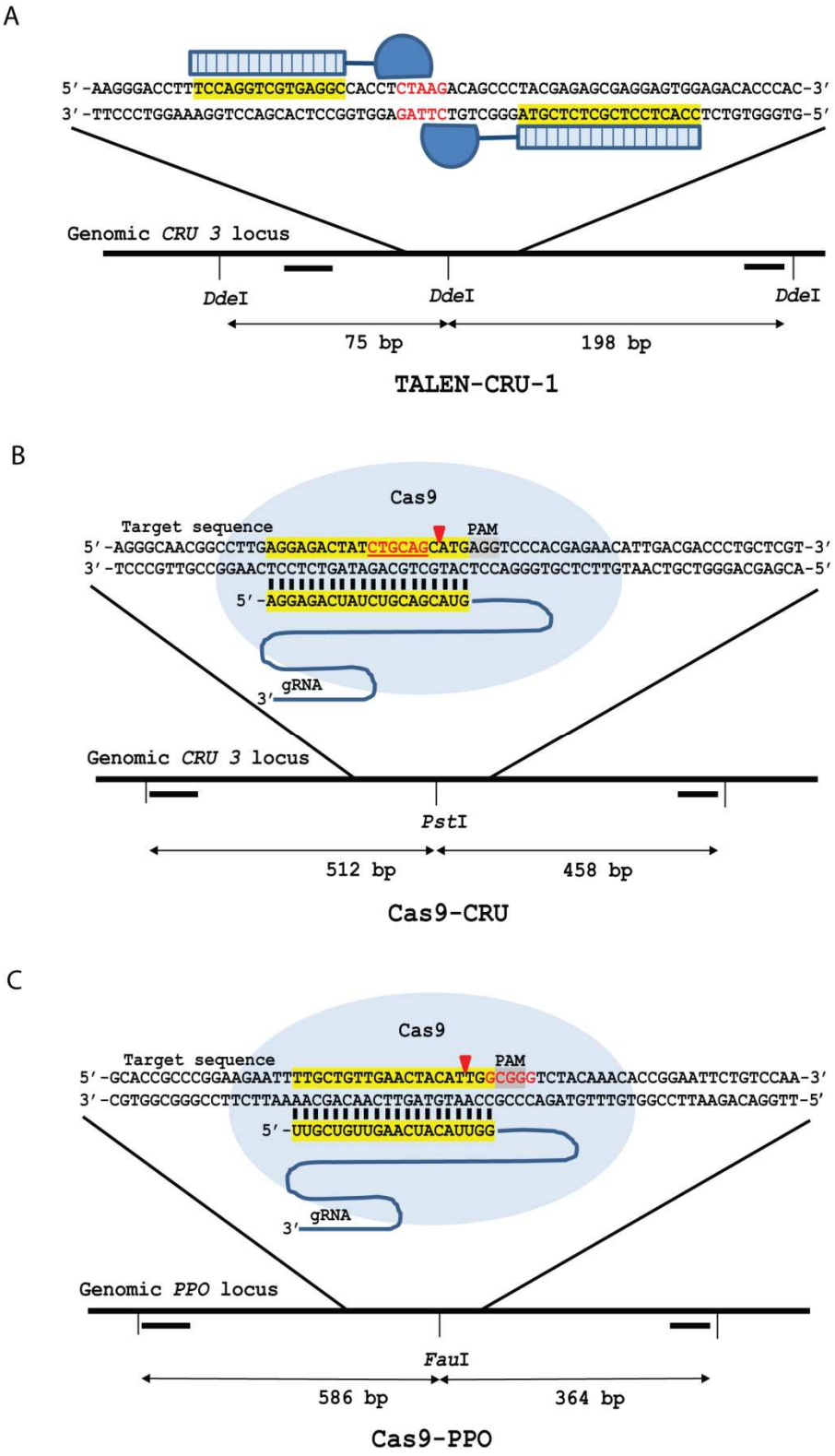


Figure 2. TALEN and CRISPR/Cas9 endonucleases for DSB induction in *CRU3* and *PPO*. TALEN-CRU-1 (**A**) and Cas9-CRU (**B**) with their target locus in the *CRU3* locus and Cas9-PPO (**C**) with its target in the PPO locus are shown. TALEN and sgRNA DNA binding sequences are highlighted with yellow, the PAM sequence is highlighted with gray and the *Dde*I, *Pst*I and *Fau*I restriction sites are shown in red lettering. The primers (**—**) used to amplify the target regions and the sizes are indicated. Red arrows indicated the position of DSB induction.

In our experimental design for both Cas9-CRU and Cas9-PPO, the *Pst*I and *Fau*I restriction sites are nearby but do not overlap the DSB site. Therefore, the preselection of loss of a restriction site has the caveat of neglecting mutations that occur outside of the restriction site. To get a more precise insight into the DSB-mediated mutations in *CRU3* and *PPO*, we used qPCR followed by high resolution melting (HRM) for footprint analysis. Then indeed, also footprints outside of the restriction site were detected (**Figure 4**). For Cas9-CRU, HRM was performed on 142 PCR clones, which revealed one deletion of 32 bp outside of the *Pst*I site. No footprints inside the *Pst*I site were detected in the remaining 141 clones, indicating a very low mutation frequency in in this plant line. For Cas9-PPO, HRM was performed on 48 PCR clones from T2 seedlings of line Cas9-PPO #7, which has a severe phenotype (**Figure 1**). Four different footprints ranging from 1 bp insertion to 5 bp deletions were found outside the *Fau*I site. None of the 48 clones contained wild-type sequences, indicating a high mutation rate in this plant line.

Taken together, these results show that our constructs of TALENS and CRISPR/Cas9 are able to induce mutations at the target sites and that our CRISPR/Cas9 constructs are more efficient than our TALENs. Only four mutations outside the restriction site were detected by the HRM method, indicating that the restriction enzyme assay gives a good estimate of the mutation frequency at our target sequences.

Increased DSB-mediated mutagenesis by CRISPR/Cas9 in c-NHEJ deficient mutants

To compare mutagenesis in wild type and NHEJ-compromised genetic backgrounds, T-DNA insertion lines *ku80*, *parp1parp2* (*p1p2*), and *ku80parp1parp2* (*ku80p1p2*) as described previously (Jia *et al.* 2012; Jia *et al.* 2013) were transformed with TALEN-CRU-1, Cas9-CRU and Cas9-PPO nucleases, and several independent primary transformants were obtained. The target region was PCR amplified using total genomic DNA from several T2 plant lines as a template, followed by restriction enzyme digestion of the PCR product. The relative band intensity was measured to estimate the mutation frequency. The mutation frequency in the TALEN-CRU-1 and Cas9-CRU lines was very low and therefore we did not perform this semi-quantification with these lines. Clear resistant bands were, however, observed in Cas9-PPO lines (**Figure 3C**). Interestingly, the mean values of the mutation frequencies in *ku80* and *ku80p1p2* lines were higher than those in the wild-type and *p1p2* lines. These results indicated that knockout of the c-NHEJ repair pathway increases DSB-mediated mutagenesis by CRISPR/Cas9. Plants appeared normal. Apparently, DSB repair of the induced DSBs was still efficient (but less precise) even in the triple mutant *ku80p1p2*.

Larger deletions are predominant in c-NHEJ deficient mutants

To assess the outcomes of DSB repair at the nucleotide level in wild type and mutant lines, genomic DNA was pre-digested and the resistant bands were purified, cloned and sequenced as described for the wild type. The results showed that there were deletions, insertions and substitutions at the *CRU3* and *PPO* target sites in mutant lines (**Figure 7**). The majority of mutations recovered in the mutant lines were deletions. Substitutions seem to be very rare events based on the sequenced data and these might be PCR artefacts.

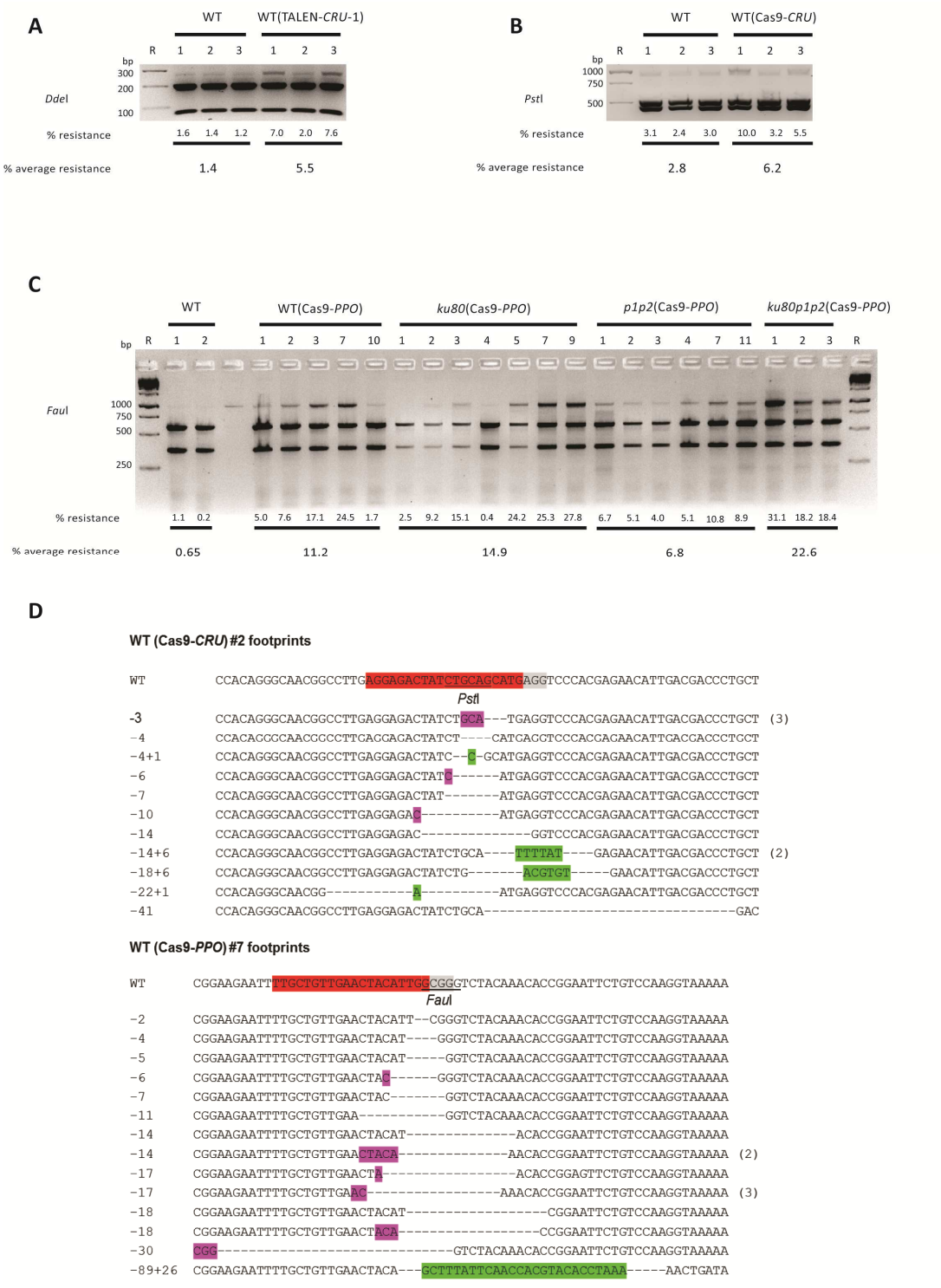
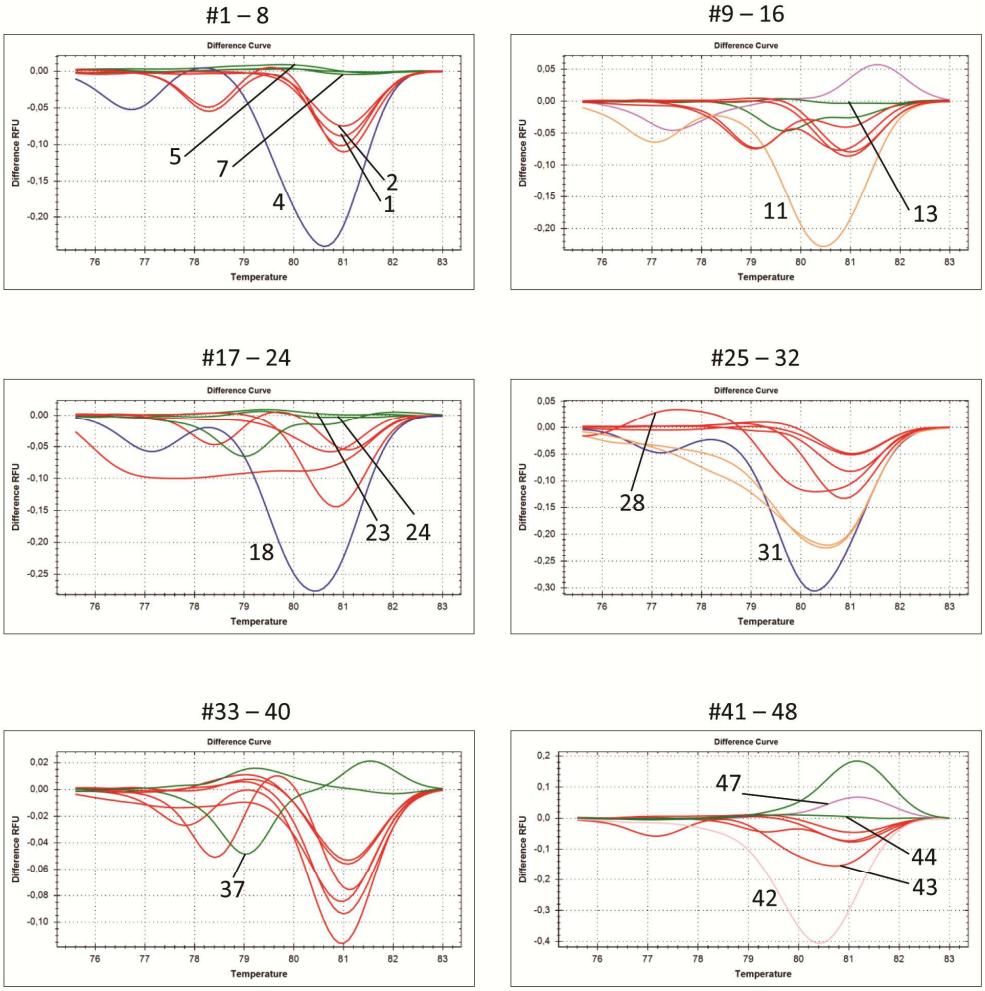


Figure 3. TALEN and CRISPR/Cas9 endonucleases induced mutagenesis. The *CRU3* target site was amplified using undigested genomic DNA from (**A**) untransformed wild type seedlings and TALEN-CRU-1 T2 transformants and digested with *Dde*I or (**B**) untransformed wild type seedlings and Cas9-CRU T2 transformants and digested with *Pst*I. (**C**) The *PPO* target site was amplified from untransformed wild type seedlings and Cas9-PPO T2 transformants of wild type and *ku80*, *parp1parp2*, and *ku80parp1parp2* mutant plant lines and digested with *Fau*I. R is the size reference (1 kb ladder) and the % resistant bands is shown below the lanes. (**D**) Sequences of *CRU3* and *PPO* targets from Cas9-*CRU* transformant #2 and Cas9-*PPO* transformant #7. The sgRNA protospacer is in red, PAM sequence is in grey, deletions are shown by dashes, insertions are in green, microhomologies used for repair are in purple. Number of multiple clones with the same sequence are shown at the right. Numbers are length of deletions (-) and insertions (+).

A



B

```
WT      GGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAAACACCGGAATTCTGTCCAAGGTAAAAAACAGC
                                    FauI
1       GGAAGAATTTTGCTGTTGAACTACATTTGGCGGGTCTACAAACACCGGAATTCTGTCCAAGGTAAAAAACAG +1
2       GGAAGAATTTTGCTGTTGAACTACATTTGGCGGGTCTACAAACACCGGAATTCTGTCCAAGGTAAAAAACAG +1
4       GGAAGAATTTTGCTGTTGAAC----------------AAACACCGGAATTCTGTCCAAGGTAAAAAACAGC -17
5       GGAAGAATTTTGCTGTTGAACTACAT-GGCGGGTCTACAAACACCGGAATTCTGTCCAAGGTAAAAAACAGC -1
7       GGAAGAATTTTGCTGTTGAACTACAT-GGCGGGTCTACAAACACCGGAATTCTGTCCAAGGTAAAAAACAGC -1
11      GGAAGAATTTTGCTGTTGAACTACG----------AAACACCGGAATTCTGTCCAAGGTAAAAAACAGC -14+1
13      GGAAGAATTTTGCTGTTGAACTACAT-GGCGGGTCTACAAACACCGGAATTCTGTCCAAGGTAAAAAACAGC -1
18      GGAAGAATTTTGCTGTTGAACTACATTG--------------ACCGGAATTCTGTCCAAGGTAAAAAACAGC -14
23      GGAAGAATTTTGCTGTTGAACTACAT-GGCGGGTCTACAAACACCGGAATTCTGTCCAAGGTAAAAAACAGC -1
24      GGAAGAATTTTGCTGTTGAACTACAT-GGCGGGTCTACAAACACCGGAATTCTGTCCAAGGTAAAAAACAGC -1
28      GGAAGAATTTTGCTGTTGAACTACAT-GGCGGGTCTACAAACACCGGAATTCTGTCCAAGGTAAAAAACAGC -1
31      GGAAGAATTTTGCTGTTGAACTACA--------------AACACCGGAATTCTGTCCAAGGTAAAAAACAGC -14
37      GGAAGAATTTTGCTGTTGAACT-----GGCGGGTCTACAAACACCGGAATTCTGTCCAAGGTAAAAGACAGC -5
42      GGAAGAATTTTGCTGTTGAAC----------------AAACACCGGAATTCTGTCCAAGGTAAAAAACAGC -17
43      GGAAGAATTTTGCTGTTGAACTACA---------CTACAAACACCGGAATTCTGTCCAAGGTAAAAGACAGC -9
44      GGAAGAATTTTGCTGTTGAACTACAT-GGCGGGTCTACAAACACCGGAATTCTGTCCAAGGTAAAAAACAGC -1
47      GGAAGAATTTTGCTGTTGAACT----TGGCGGGTCTACAAACACCGGAATTCTGTCCAAGGTAAAAAACAGC -4
```

Figure 4. HRM analysis of the PPO target. HRM analyses were performed on 48 PCR clones from pool of 10 seedlings of wild-type T2 seedlings of Cas9-*PPO* transformant #7. (**A**) Difference melt curves of samples 1 – 48 measured in relative fluorescence units (RFU). Numbers indicated in the graph refer to the sequences below. (**B**) Sequences of representative *PPO* targets. *PPO* sgRNA protospacer (red), the PAM sequence (gray) and *Fau*I restriction site (underlined) are indicated in the WT sequence. Footprints included deletions (dashed lines), insertions (green) and substitutions (blue). Microhomologies used for repair are shown in purple.

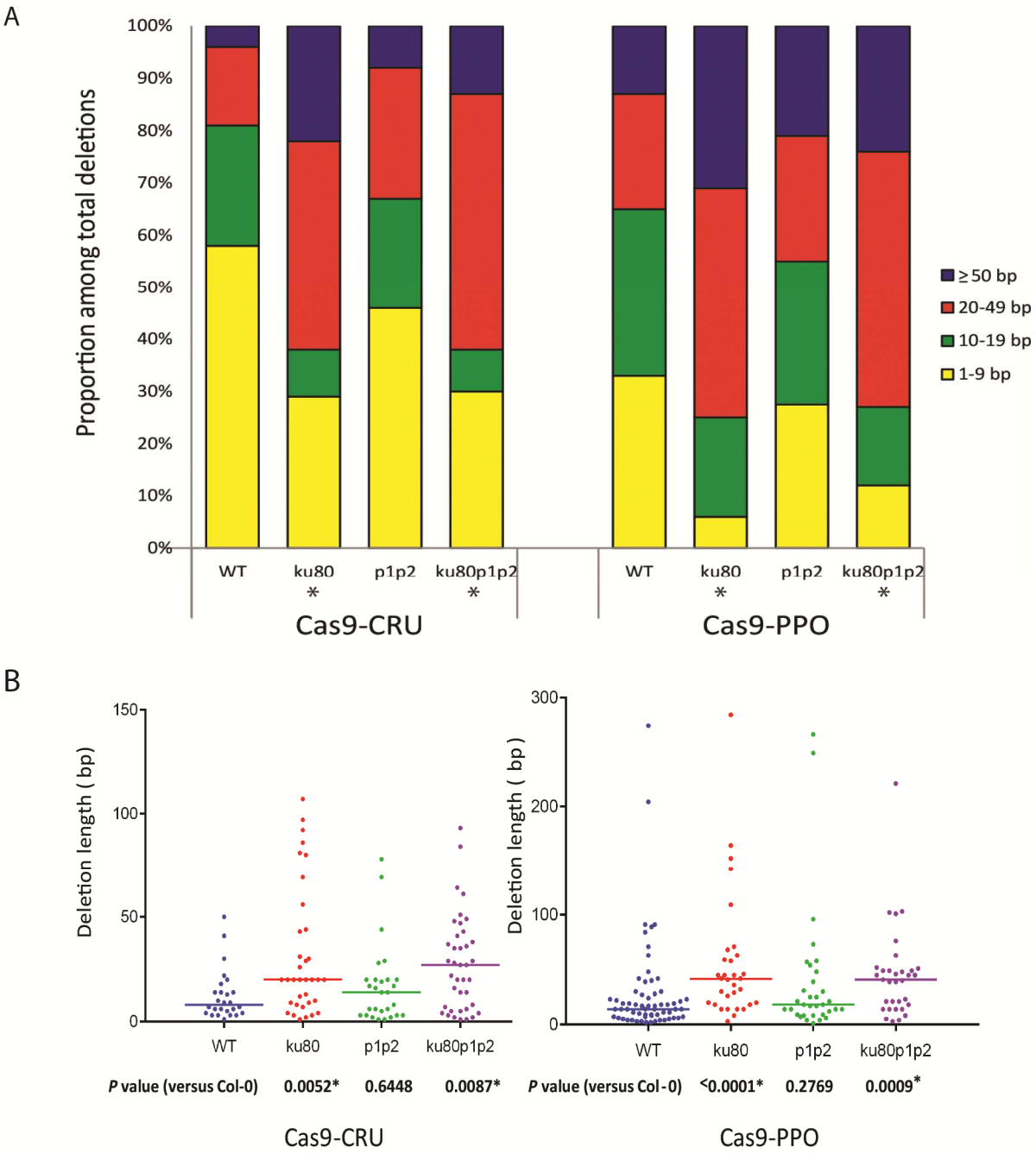


Figure 5. Analysis of deletion length. (**A**) Distribution of deletion lengths for the indicated genotypes with Cas9 nucleases. (**B**) Scatter plot of deletion lengths for the indicated genotypes. Median deletion length is indicated at the bar on the graph. P-values are derived from a two-tailed Mann-Whitney test. Asterisk indicates a statistically significant difference from wild-type ($P<0.05$).

Due to the small size of the amplified target region in the TALEN-CRU-1 lines no large deletions could be found. Therefore we did not include the TALEN-CRU-1 lines in the analysis of deletion length and we examined the length of deletions only from all genotypes expressing Cas9-CRU or Cas9-PPO (**Table 3**, **Figure 5A**). In wild-type Cas9-CRU transformants, 57% of deletions were <10 bp and about 23% ranged from 10 to 19 bp, 15% ranged from 20 to 49 bp and 5% were ≥50 bp. The *p1p2* mutant lines showed somewhat

longer deletions; about 25% of deletions ranged from 20 to 49 bp and 7% were ≥50 bp. Larger deletions were predominant in the *ku80* and *ku80p1p2* mutant lines. In *ku80* lines 62% of the deletions were larger than 20 bp (22% were ≥50 bp), and in the *ku80p1p2* lines 61% of the deletions were larger than 20 bp (12% were ≥50 bp).

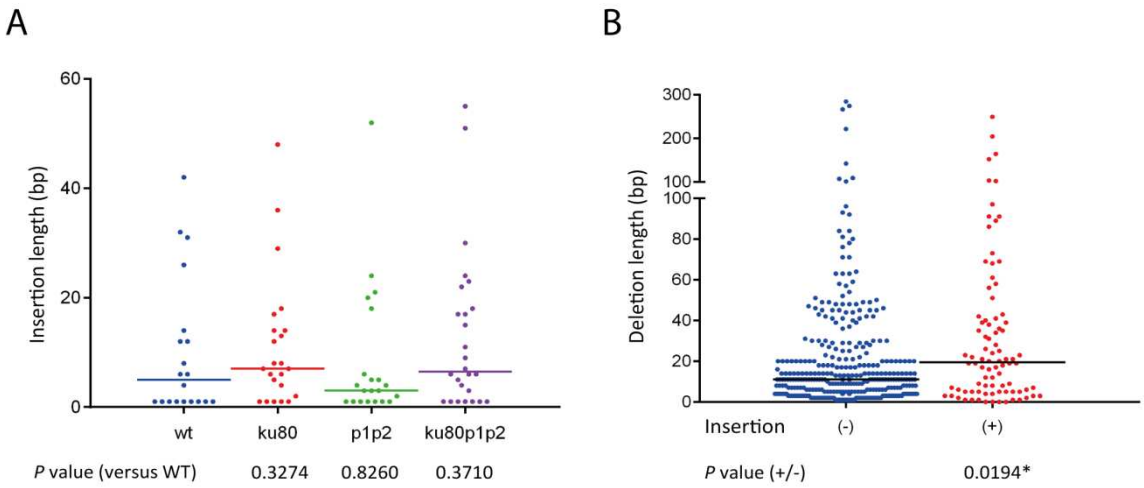
Deletion lengths in Cas9-PPO transformants were also examined (**Figure 5**). Similar to Cas9-CRU, there were no big differences in deletion length between the wild type and *p1p2* mutant lines. In the wild type about 33% of deletions were <10 bp, 32% ranged from 10 to 19 bp, 22% ranged from 20 to 49 bp and 13% were ≥50 bp. In *p1p2* lines 27% of deletions were <10 bp, 27% ranged from 10 to 19 bp, 24% ranged from 20 to 49 bp and 21% were ≥50 bp. Larger deletions of the *PPO* target were however, again predominant in *ku80* and *ku80p1p2* mutant lines. About 75% of deletions in *ku80* lines were larger than 20 bp, and about 73% of deletions in *ku80p1p2* lines were larger than 20 bp.

We performed statistical analysis using a two-tailed Mann-Whitney test, to find out whether the observed differences were significant. For the Cas9-CRU and Cas9-PPO nucleases, comparison of deletion lengths in wild type to *ku80* and *ku80p1p2* lines showed a statistically significant difference, whereas comparison of deletion lengths in wild type to *p1p2* lines did not (**Figure 5B**). These results indicate that imprecise end-joining after loss of the c-NHEJ key component KU80 resulted in substantial increases in deletion length and suggests a shift to a more error-prone repair mechanism of DSB repair in absence of KU80.

Templated insertions in wild type and NHEJ mutants

Insertion events, sometimes accompanied by deletions, were observed at the target loci in TALEN-CRU-1, Cas9-CRU or Cas9-PPO transformed wild type and mutant lines, although less frequently than deletions. The insertion frequency in mutants was comparable to or higher than the insertion frequency in the wild type. More than half of the insertions were smaller than 10 bp. A maximum insertion length of 60 bp was observed. Furthermore, insertion lengths in NHEJ mutant lines were not significantly different from wild type when insertion data of all nucleases was combined, indicating that the insertion mechanism may be independent of KU80 and PARP (**Figure 6A**). In addition, from the combined insertion data of all nucleases it can be deduced that the deletion length of junctions with insertions were significantly larger than junctions without insertions (**Figure 6B**).

Interestingly, many inserted sequences have at least one match to DNA within 100 bp of the repaired DSB. Some insertions have complex compositions with multiple stretches of identity, including reverse complementary homology. These results suggest that polymerase θ may be involved in the repair of these DSBs (Roerink *et al.* 2014). Another signature of Pol θ-mediated DSB repair is the presence of sequence identity between the 3' end that generated the junction (the primer) and the sequence immediately upstream of the template that is used for DNA synthesis. Such sequence identity is present in about 50 % of the inserted sequences (**Figure 6C**). The *ku80* and *ku80p1p2* mutant lines appeared to have more templated-insertion events than wild type and *p1p2* lines, although such insertions were found in all four genotypes (**Table 4**). Therefore, the templated insertions probably resulted from a KU80- and PARPs-independent alternative end-joining mechanism, such as that mediated by the recently discovered Pol θ (van Kregten *et al.* 2016).



C

Templated insertions:

1. *ku80* TALEN-CRU-1	-5+6	AAGGGACCTTTCCAGGTCGTGAGGCCACCTCTAGACACCGCCCTACGAGAGCGAGGAGTGGAGACACCCACGTAGCCCA
2. *ku80* TALEN-CRU-1	-5+5	AAGGGACCTTTCCAGGTCGTGAGGCCACCTCTAACACCCCCTACGAGAGCGAGGAGTGGAGACACCCACGTAGCCCAC
3. *ku80* Cas9-CRU	-12+8	CAACGGCCTTGAGGAGACTATCTGCA---TCGTGCTG--------ACGAGAACATTGACGACCCTGCTCGTGCTGACGT
4. *ku80* Cas9-CRU	-97+12	CAGGGCAACGGCCTTGAGGAG-----TATGTCAGGCTC-----ATACCTTGCCCATCTTGGAGTATGTCAGGCTCAGTG
5. *ku80* Cas9-CRU	-17+10	ACAGGGCAACGGCC------ACGTATACAA----ATGAGGTCCCACGAGAACAT(210 bp)AGAACGTATACAAATGT
6. *ku80p1p2* Cas9-CRU	-41+22	CC-----TGCTCGTGCTGACGTGTACAAG----CCCACGAGAACATTGACGACCCTGCTCGTGCTGACGTGTACAAGCC
7. WT Cas9-PPO	-91+31	TTGCTCGTATTTCTTCAGGAACTATCTACAGCTCCTCA--GCTCGTATTTCTTCAGGAACTATCTACAGCT---ACAGC
8. *ku80* Cas9-PPO	-32+13	TTGCTCGTATTTCTTCAGGAACTATCTACAGCTCCTCCTCTTTCCAAATCGCGCACCGC----TCGTATTTCTTCA--- ---GGGTCTACAAA
9. *p1p2* Cas9-PPO	-73+18	CGGAA----------CTGATAGTTTTCTTCTCT----------CTTGTAACACATCTTTATTCAACCAAGTAAAC CTAAGAACTGATAGTTTTCTTCTCTCTCT
10.*Ku80p1p2* Cas9-PPO	-103+15	GTATTTCTTCAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCGCGCACCGCCCGGAAGAATTTTGCTGTTGAACT- ---TTTCTTCAGGAACTA----TCTT
11.*ku80* TALEN-CRU-1	-7+7	AAGGGACCTTTCCAGGTCGTGAGGCCACCTCTTTCACCCCCCTACGAGAGCGAGGAGTGGAGACACCCACGTAGCCCAC
12.*ku80* Cas9-CRU	-20+14	GGGCAACGGCCTTGAGGAGACTATCTG-ACGTGTTGACGTGT-ACATTGACGACCCTGCTCGTGCTGACGTGTACAAGC
13.*ku80* Cas9-CRU	-86+8	GGGCAAC----------ACCATCGT----------TCGCGTGACCAGCGTCAACAG
14.*ku80* Cas9-CRU	-16+21	ACCCACGTAGCCCACAGGGCAACGGCCTTGAGTGGACCACAGGGCAACGGCCTTGAGGTCCCACGTGAACATTGACCCT
15.*ku80p1p2* Cas9-CRU	-28+17	GCGAGGAGTGGAGACACCCACGTAGCCCACAGG-------AGTGGAGACACCCACCT----CATGAGGTCCCACGAGAA
16.WT Cas9-PPO	-24+12	CGGAAGAATTTTGCTGTTGAA---TTCTGTCCGCTT---------CGGAATTCTGTCCAAGGTAAAAA
17.*Ku80p1p2* Cas9-PPO	-40+9	TTTCCAAATCGCGCACCGCCCGGAAGAATTTTGCTGTTGAACTACA---GCAAACAGC---AAACAGCAAACACTTGTA
18.*ku80p1p2* Cas9-CRU	-61+24	CCACAGGGCAACGGCC----------CACGAGCAGGGTCGTCAATGTTCT----------CGTGTA
19.*ku80p1p2* Cas9-CRU	-43+18	ACCCACGTA-------CCCAACAGGGTCGTCAAT--------GTCCCACGAGAACATTGACGACCCTGCTCGTGCTGAC
20.*ku80p1p2* Cas9-CRU	-35+11	CCACAGG----------TCAATGTTATA----------TCCCACGAGAACATTGACGACCCTGCTCGTGCTGAC
21.WT Cas9-PPO	-23+42	TCCTCACTCTTTCCAAATCGCGCACCGCCCGGAAGAATTTTGCTGTTGAACATTTGGAAAGAGTGCGCAAACACTTGTA ACACATCAAAACACCGGAATTCTGTCCAAGGTAAAAAACAGCAAACACTTGTAACACATCTT
22.WT Cas9-PPO	-8+32	TTTCTTCAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCGCGCACCGCCCGGAAGAATTTTGCTGTTGAACTA--- --TCTACAGCTTCAGGAACTATCTACAGCGCC---CGGGTCTACAA
23.*Ku80p1p2* Cas9-PPO	-5+30	CAAATCGCGCACCGCCCGGAAGAATTTTGCTGTTGAACTACAGCAAAATTCTGTTGAATTTTTCTGTTGCTTGGGTCTA
24.*Ku80p1p2* Cas9-PPO	-3+51	TTTCCAAATCGCGCACCGCCCGGAAGAATTTTGCTGTTGAACTACATCTTTATTCTTCCGGGCGGTGCGCGATTTGATA TCTCAAATCGCGCACCGCCCGGGTC

NNNN Insertion NNNN Template NNNN Inverse Template NNNN Homology Nucleotides

Figure 6. Analysis of insertions. **(A)** Scatter plot of insertion lengths for the indicated genotypes. Data are combined from all nucleases. **(B)** Scatter plot of deletion lengths with or without insertions. Data are combined from all genotypes and nucleases. Median insertion or deletion length is indicated by the bar in the graph. P-values are derived from a two-tailed Mann-Whitney test. Asterisk indicates the statistical significant difference ($P<0.05$). **(C)** Footprints consisting of deletions (dashes) accompanied with insertions. Insertions are shown in green, template sources for the insertions are shown in yellow (direct strand) or underlined (reverse complement). Homologies between sequences flanking the template and the insertion and probably used as primer are shown in gray. Footprints from 1 to 10 are examples of perfectly matching the template, 11 to 17 are partially matching the template and 18 to 24 are reversely matching the template. Numbers are length of deletions (-) and insertions (+).

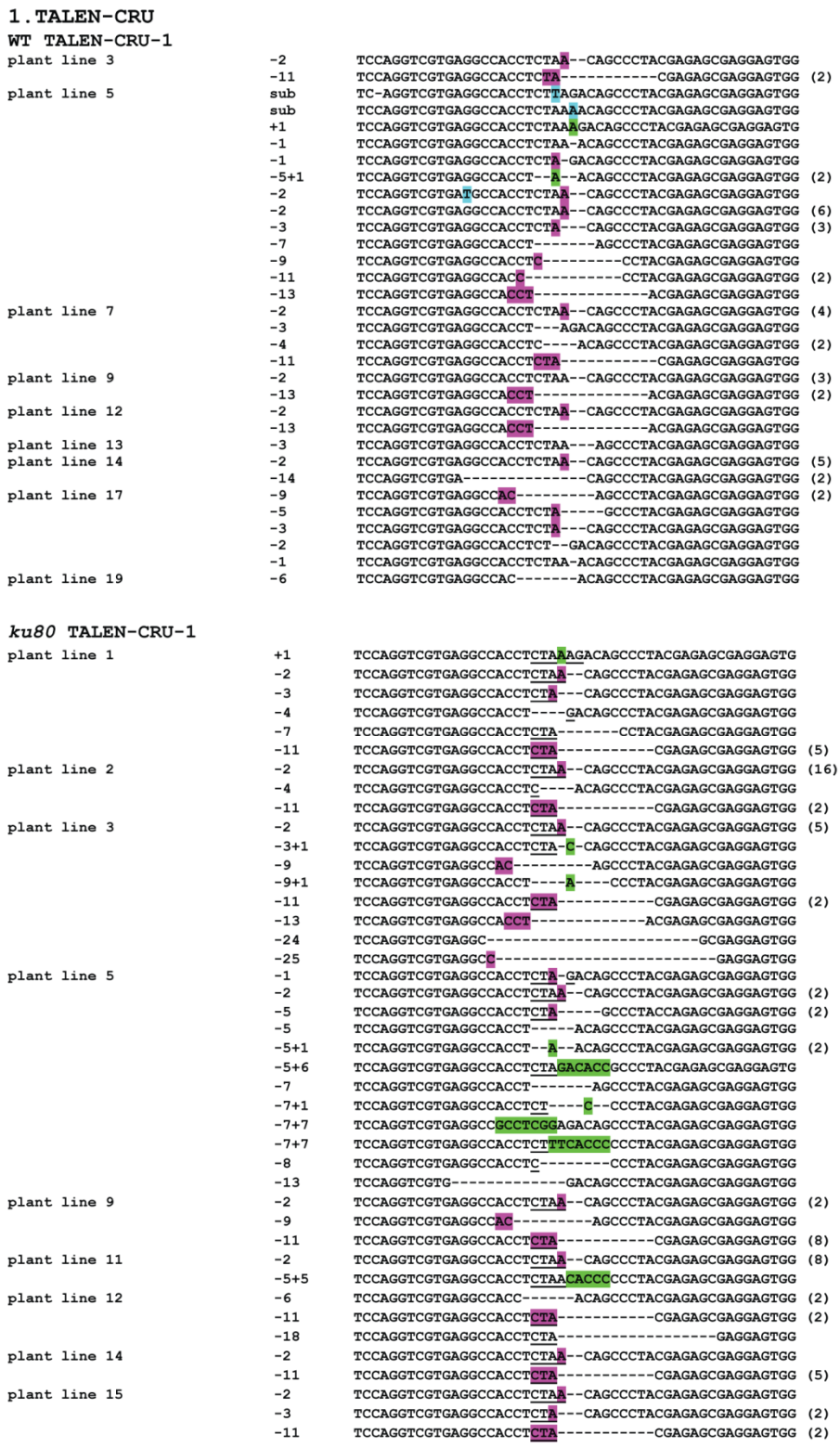


Figure 7. Sequences of resistant target sites. DNA from several plant lines of wild type and NHEJ mutants with different nuclease constructs was predigested with the appropriate restriction enzyme (*Dde*I for TALEN-CRU-1, *Pst*I for Cas9-CRU, and *Fau*I for Cas9-PPO), used for PCR, digested with the same enzyme and resistant products were cloned and sequenced. Footprints included deletions (dashed lines), insertions (green) and substitutions (blue). Microhomologies used for repair are shown in purple. Numbers are length of deletions (-) and insertions (+).

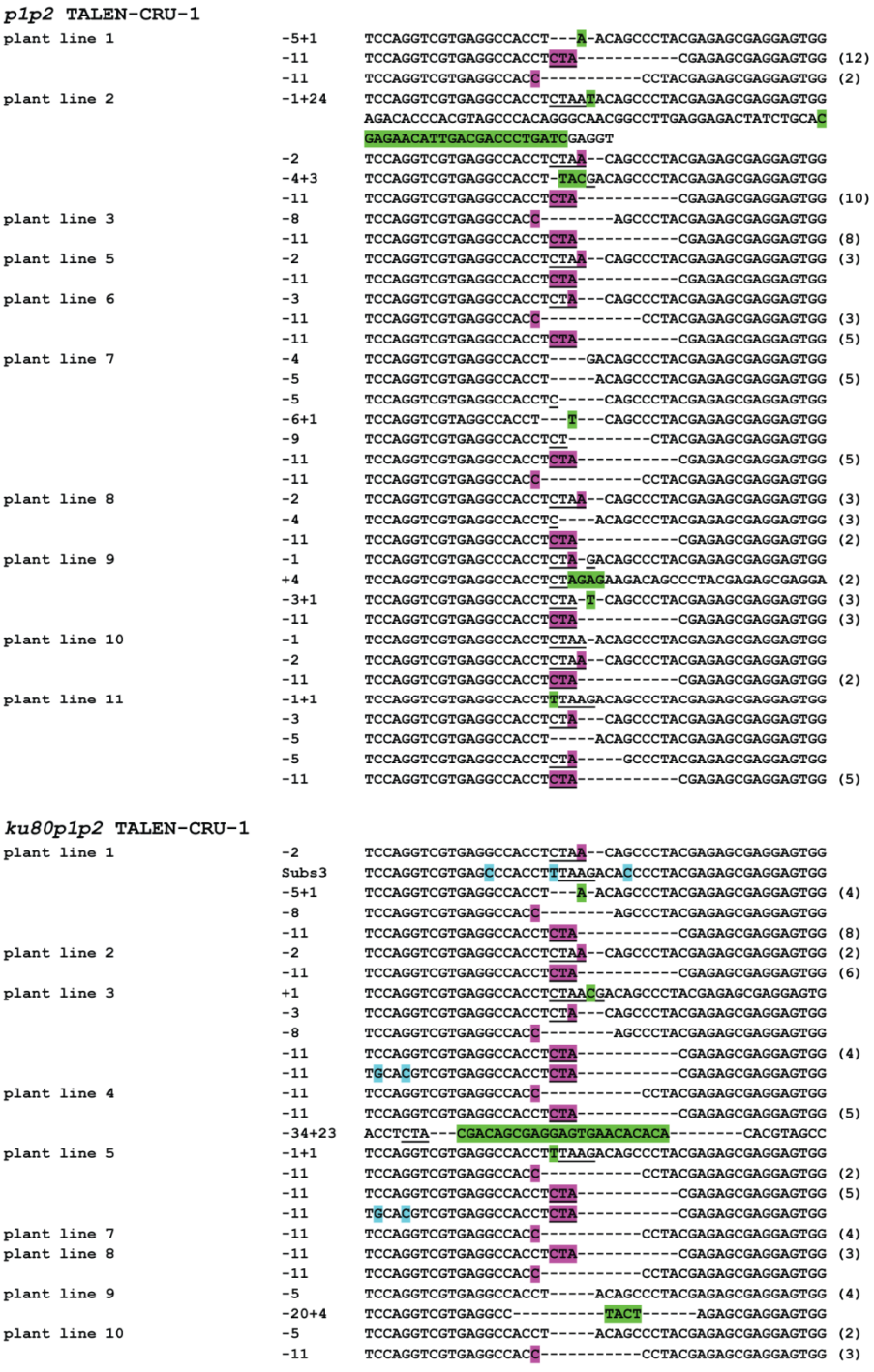


Figure 7. (continued 1).

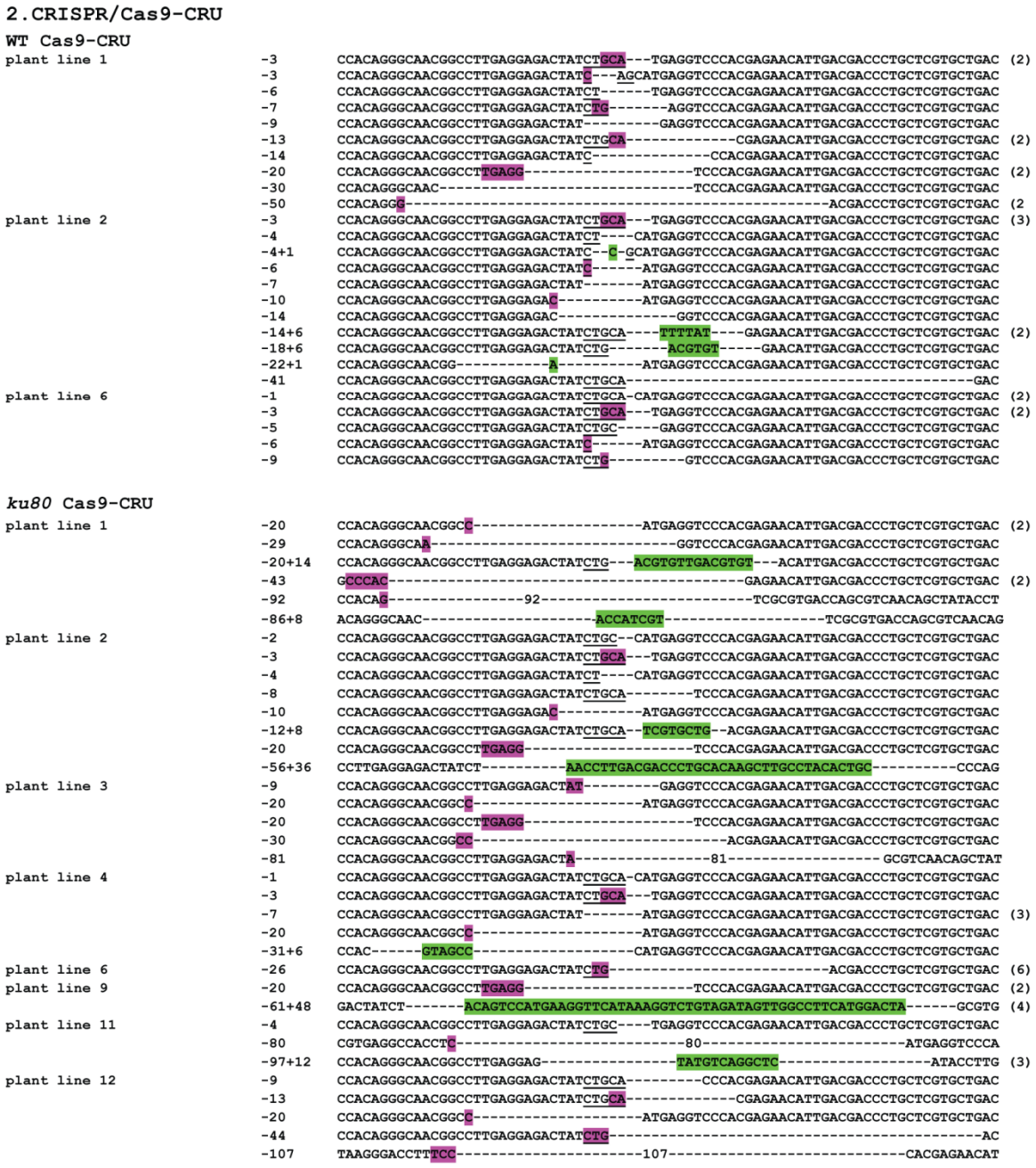


Figure 7. (continued 2).

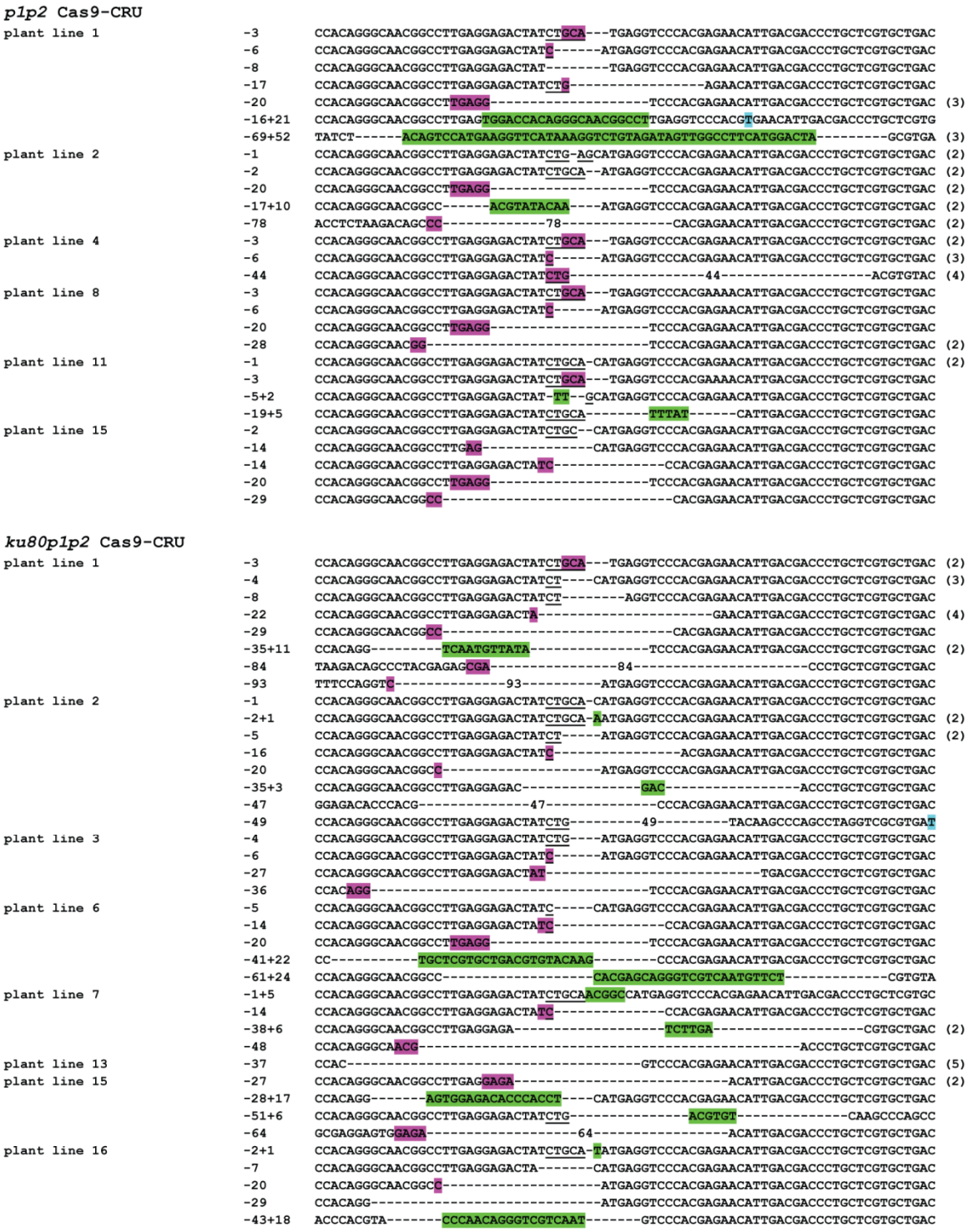


Figure 7. (continued 3).

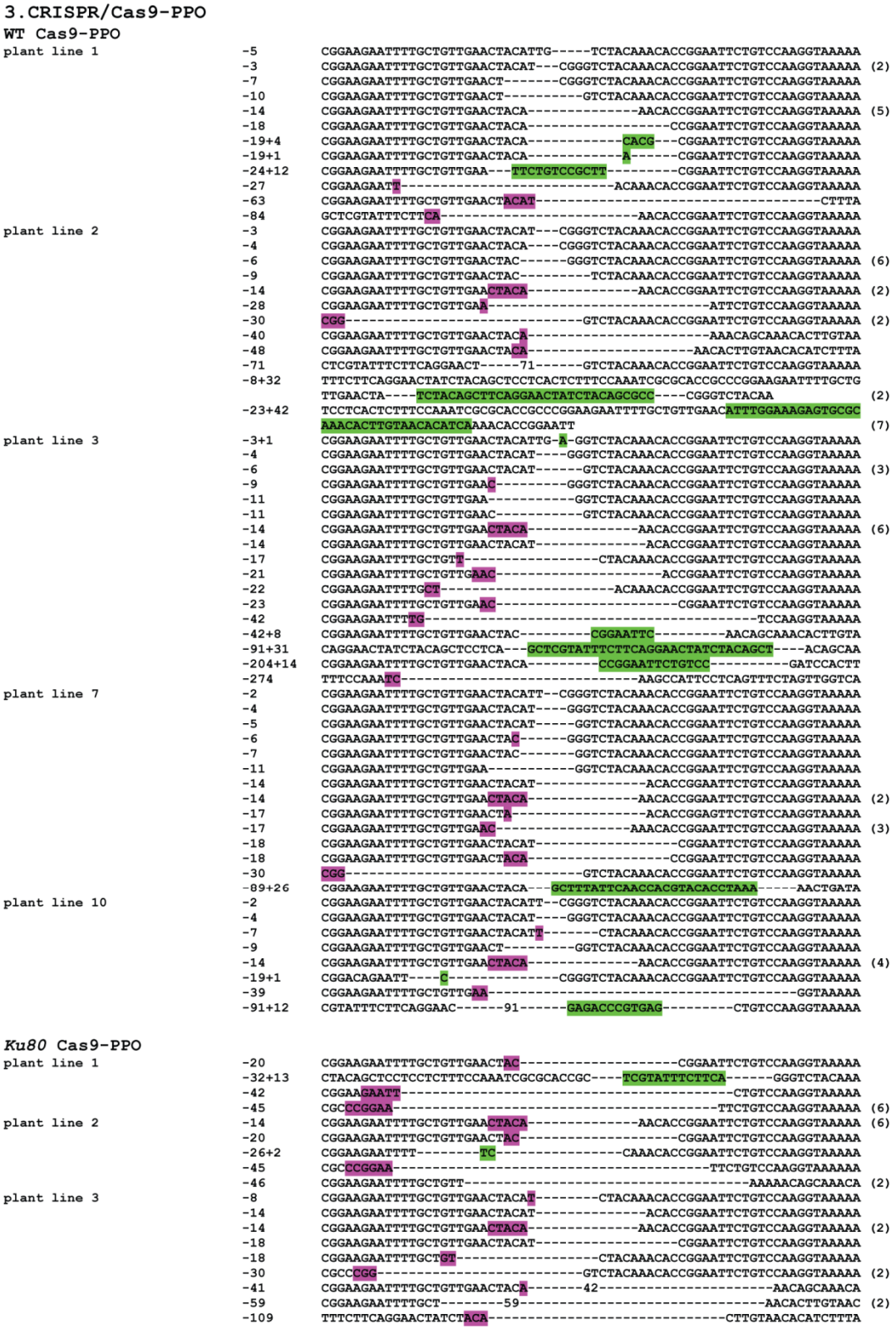


Figure 7. (continued 4).

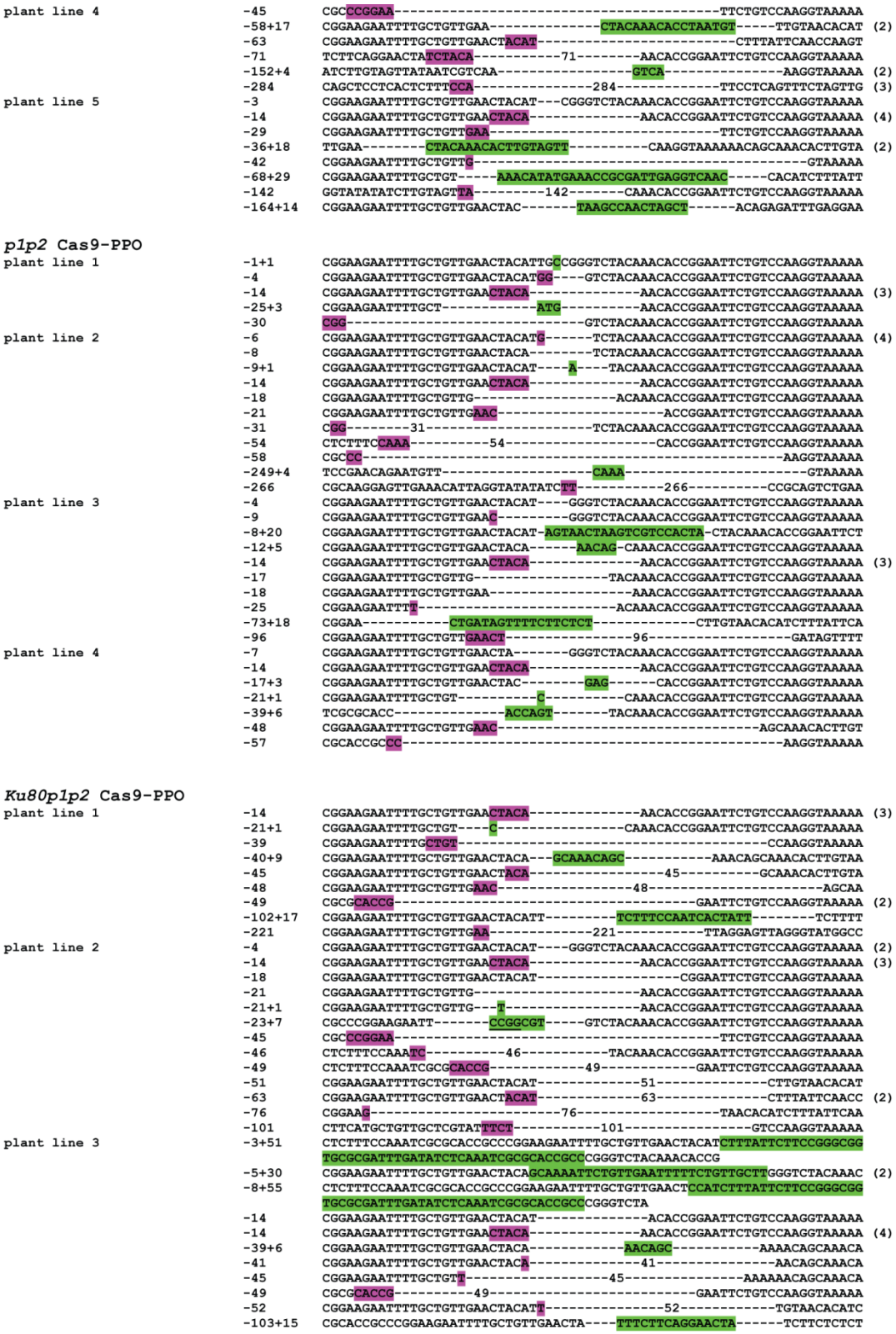


Figure 7. (continued 5).

Discussion

In this study, we demonstrated that TALEN- and CRISPR/Cas9-induced DSBs can be successfully repaired in *Arabidopsis*, even after loss of key components of the NHEJ repair pathway. In general plants expressing these nucleases look healthy and develop normally (unless an essential gene was targeted). As precise repair restores the break site leading to an intact substrate for the nuclease, a cycle of DSB induction and repair continues until mutations in the target sequence prevent the action of the nuclease. We showed that especially the Cas9-PPO nuclease was a very efficient tool for targeted mutagenesis, with close to 100% mutation frequency in one line. Since the *PPO* gene is an essential gene, this resulted in stunted growth of the seedlings (**Figure 1**). This phenotype was not observed in earlier targeted mutagenesis experiments of *PPO* using ZFNs (de Pater *et al.* 2013), indicating a higher activity of CRISPR/Cas9 on the *PPO* gene compared to the ZFNs. We noticed that the sgRNA recognized a sequence in the *PPO* gene with GG just 5'of the PAM sequence, which was recently shown to promote higher rates of mutagenesis (Farboud and Meyer 2015).

We analysed TALEN-induced and Cas9-induced mutations. Both nucleases are able to induce mutagenesis at the target sites and a variety of footprints (deletions and insertions) were found. However, much longer deletions were found in Cas9-CRU lines than in TALEN-CRU-1 lines with the same genetic background. The main reason for this observation may be the different lengths of amplified target sequence. In TALEN-CRU-1 lines a region of only 273 bp was amplified, due to the presence of additional *Dde*I sites just outside of this region. In Cas9-CRU lines a region of 970 bp long was amplified and analysed for footprints. This means that larger deletions could be missed more easily in TALEN-CRU-1 lines than in Cas9-CRU lines. Besides, different types of break ends may activate different end-resection mechanisms, which can affect end-joining outcomes. TALENs induce DSBs with 5' overhangs which are suitable for ligation, whereas Cas9 induces mostly blunt ends or incompatible ends (Jinek *et al.* 2012). Additionally, CRISPR-Cas9 might be more efficient than TALENs in plants.

As a result of imperfect end joining, various mutations in the target sites were found in each line. In the analysis of DSB repair outcome in NHEJ mutants, we observed a statistically significant increase in the median deletion length at the repair junction in the *ku80* and *ku80parp1parp2* mutants compared to wild type, but not in the *parp1parp2* mutant. This suggests that b-NHEJ is a more error-prone DSBs repair pathway than c-NHEJ. KU is known to competitively bind to DSB ends and protects break ends from end processing (Downs and Jackson 2004; Fell and Schild-Poulter 2014). Thus, when KU is absent, DNA ends are exposed to end resection proteins which would promote the generation of larger deletions. Similar results have been described previously with ZFNs-induced DSB repair in a *ku80* mutant and in *ku70* and *lig4* mutants (Osakabe *et al.* 2010; Qi *et al.* 2013)

We previously showed that PARP1 and PARP2 are involved in the MMEJ repair pathway by an *in vitro* end joining assay in *Arabidopsis* (Jia *et al.* 2013). In the *in vivo* end-joining experiments described here, however, we did not observe a role for PARP1 and PARP2 in MMEJ, and therefore there must be another repair pathway independent of PARP1 and PARP2 that uses microhomology. It is still elusive whether b-NHEJ is a single pathway or a category containing multiple mechanisms (Deriano and Roth 2013). The similar mutation

characteristics observed in the *parp1parp2* mutant and the wild type supports the dominant role of the KU-dependent c-NHEJ pathway rather than the PARP-dependent b-NHEJ pathway. Notably, we did not observe much differences in junction characteristics between *ku80* and *ku80parp1parp2* mutants, indicating other repair pathways (independent of PARP) with similar characteristics become active when c-NHEJ is absent. However, we cannot rule out that other factors, for example PARP3 (Rulten *et al.* 2011; Langelier *et al.* 2014), could slip into the b-NHEJ pathway without disturbing its outcome.

Insertions were found at both c-NHEJ-proficient and -deficient repair junctions, although most junctions were repaired without an insertion. *ku80* mutants had more insertion events than the wild type. However, the median insertion length in break junctions is a few base pairs and no statistically significant difference was observed among mutants and wild type. Besides, larger deletions (of more than 20 bp) were found with insertions compared to those in repair products without insertions during Cas9-induced repair. Templated insertions were observed both in c-NHEJ efficient and c-NHEJ deficient mutants in animal cells. The current models of templated mutagenesis are based on a MMEJ mechanism involving DNA Polymerase θ (McVey and Lee 2008; Koole *et al.* 2014; Roerink *et al.* 2014). In plants, templated insertions were also observed after DSB induced repair in *Arabidopsis* and tobacco (Shirley *et al.* 1992; Gorbunova and Levy 1997; Salomon and Puchta 1998; Lloyd *et al.* 2012; Vu *et al.* 2014). Furthermore, a recent study showed that the *Arabidopsis* Pol θ ortholog *Tebichi* (*Teb*) is essential for T-DNA integration (van Kregten *et al.* 2016). Templated insertions were found at the repair junctions of T-DNA inserts, and it was shown that *teb* mutants were resistant to T-DNA integration and very sensitive to the DNA damaging agents bleomycin and MMS. Our results indicate that nuclease-induced DSBs may be repaired by Ku-dependent NHEJ, or a backup pathway in the absence of Ku, leading to larger deletions in the latter case. Templated insertions, which have the hallmarks of Pol θ-mediated repair, may be formed in both cases, but in with a higher frequency in the absence of Ku, revealing a complex interplay of repair factors during DSB repair in *Arabidopsis*.

Acknowledgements

We thank the Puchta lab for supplying the pEn-Chimera and pDe-Cas9 vectors. We also like to thank BSc internship students Manoah van de Velde and Kübra Kontbay, and MSc internship student Kelly van Kooperen for their technical assistance. This work was financially supported by the Partnership Program STW-Rijk Zwaan of the Dutch Technology Foundation STW, which is part of the Netherlands Organisation for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs, Agriculture and Innovation (12428). HS was supported by a scholarship from the Chinese Scholarship Council (CSC) and PJJH by the fellowship associated with the appointment as KNAW Academy Professor.

Table 1. TALEN RVDs

TAL	RVDs
TAL-CRU-1-left	NI HD HD NI NN NN NG HD NN NG NN NI NN NN **HD**
TAL-CRU-1-right	HD HD NI HD NG HD HD NG HD NN HD NG HD NG HD NN NG **NI**

Table 2. Primers used for cloning and PCR reactions.

Primer	Sequence	Used for
SP509	ATTGAGGAGACTATCTGCAGCATG	sgRNA cloning CRU3
SP510	AAACCATGCTGCAGATAGTCTCCT	sgRNA cloning CRU3
SP512	ATTGTTGCTGTTGAACTACATTGG	sgRNA cloning PPO
SP513	AAACCCAATGTAGTTCAACAGCAA	sgRNA cloning PPO
SP491	GCTTCAGAACCAACAAGACAGC	*CRU3* target sense (TALEN)
SP492	TGAGCCTGACATACTCCAAG	*CRU3* target antisense (TALEN/HRM)
SP245	TGCCAACACTCCAGGCTCTG	*CRU3* target sense (CRISPR/Cas9)
SP248	CAAGTGGTCAACGACAACGG	*CRU3* target antisense (CRISPR/Cas9)
SP392	CACTTTGACAGATTAGGTAG	*PPO* target sense (CRISPR/Cas9)
SP538	CTTCCACTAACTCACCTTC	*PPO* target antisense (CRISPR/Cas9)
SP560	CTCCTCACTCTTTTCCAAATCG	*PPO* target sense (HRM)
SP561	AGATGTGTTACAAGTGTTTGCTG	*PPO* target antisense (HRM)
SP563	ACCTCTAAGACAGCCCTACG	*CRU3* target sense (HRM)

Table 3. Distribution of deletion length for the target site derived from the indicated genotypes.

Nuclease mutants	Deletion length				Total deletions
	1-9 bp	10-19 bp	20-49 bp	≥50 bp	
WT Cas9-CRU	15 (57.7%)	6 (23.1%)	4 (15.4%)	1 (3.8%)	26
ku80 cas9-CRU	10 (28.6%)	3 (8.6%)	14 (40%)	8 (22.8%)	35
p1p2 Cas9-CRU	13 (46.4%)	6 (21.4%)	7 (25%)	2 (7.2%)	28
ku80p1p2 Cas9-CRU	12 (30.8%)	3 (7.7%)	19 (48.7%)	5 (12.8%)	39
WT Cas9-PPO	21 (33.3%)	20 (31.7%)	14 (22.2%)	8 (12.8%)	63
ku80 Cas9-PPO	2 (6.3%)	6 (18.7%)	14 (43.7%)	10 (31.3%)	32
p1p2 Cas9-PPO	9 (27.3%)	9 (27.3%)	8 (24.2%)	7 (21.2%)	33
ku80p1p2 Cas9-PPO	4 (12.1%)	5 (15.2%)	16 (48.5%)	8 (24.2%)	33

Table 4. Insertions and templated-insertions.

Nuclease mutants	Total mutations	Insertions	Templated-insertions >3 bp
WT TALEN-CRU-1	30	2 (6.6%)	0
ku80 TALEN-CRU-1	42	9 (21.4%)	3 (7.1%)
p1p2 TALEN-CRU-1	36	7 (19.5%)	1 (2.8%)
ku80p1p2 TALEN-CRU-1	26	5 (19.2%)	0
WT Cas9-CRU	26	4 (15.4%)	0
ku80 Cas9-CRU	35	7 (20.0%)	6 (17.1%)
p1p2 Cas9-CRU	28	5 (17.9%)	0
ku80p1p2 Cas9-CRU	39	11 (28.2%)	5 (12.8%)
WT Cas9-PPO	63	12 (19.1%)	4 (6.3%)
ku80 Cas9-PPO	32	7 (21.9%)	4 (12.5%)
p1p2 Cas9-PPO	33	10 (30.3%)	1 (3.0%)
ku80p1p2 Cas9-PPO	33	10 (30.3%)	6 (18.2%)

References

Alonso, J. M., M. Alonso, A. N. Stepanova, T. J. Leisse, D. K. Stevenson *et al.*, 2009 Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science 301: 653–657.

Cermak, T., E. Doyle, and M. Christian, 2011 Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. 39: e82.

Christian, M., T. Cermak, E. L. Doyle, C. Schmidt, F. Zhang *et al.*, 2010 Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186: 756–761.

Clough, S. J., and A. F. Bent, 1998 Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J. 16: 735–743.

Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto *et al.*, 2013 Multiplex genome engineering using CRISPR/Cas systems. Science 339: 819–823.

Curtis, M. D., and U. Grossniklaus, 2003 A Gateway cloning vector set for high-throughput functional analysis of genes in planta. Breakthr. Technol. 133: 462–469.

Deriano, L., and D. B. Roth, 2013 Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. Annu. Rev. Genet. 47: 433–455.

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

Farboud, B., and B. J. Meyer, 2015 Dramatic enhancement of genome editing by CRISPR/cas9 through improved guide RNA design. Genetics 199: 959–971.

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

Fell, V. L., and C. Schild-Poulter, 2014 The Ku heterodimer: Function in DNA repair and beyond. Mutat. Res. Mutat. Res. 763: 15-29.

Gorbunova, V., and A. A. Levy, 1997 Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. Nucleic Acids Res. 25: 4650–4657.

Jia, Q., P. Bundock, P. J. J. Hooykaas, and S. de Pater, 2012 Agrobacterium tumefaciens T-DNA Integration and Gene Targeting in Arabidopsis thaliana Non-Homologous End-Joining Mutants. J. Bot. 2012: 1–13.

Jia, Q., A. Den Dulk-Ras, H. Shen, P. J. J. Hooykaas, and S. de Pater, 2013 Poly(ADP-ribose)polymerases are involved in microhomology mediated back-up non-homologous end joining in Arabidopsis thaliana. Plant Mol. Biol. 82: 339–351.

Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna *et al.*, 2012 A Programmable Dual-RNA – Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science 337: 816–822.

Jinek, M., F. Jiang, D. W. Taylor, S. H. Sternberg, E. Kaya *et al.*, 2014 Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science 343: 1247997.

Koole, W., R. van Schendel, A. E. Karambelas, J. T. van Heteren, K. L. Okihara *et al.*, 2014 A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. Nat. Commun. 5: 3216.

van Kregten M., S. de Pater, R. Romeijn, R. van Schendel, P. Hooykaas and M. Tijsterman, 2016 T-DNA integration in plants results from Polymerase Theta-mediated DNA repair. Nature plants 2:1-6.

Langelier, M. F., A. a. Riccio, and J. M. Pascal, 2014 PARP-2 and PARP-3 are selectively activated by 5' phosphorylated DNA breaks through an allosteric regulatory mechanism shared with PARP-1. Nucleic Acids Res. 42: 7762–7775.

Lloyd, A. H., D. Wang, and J. N. Timmis, 2012 Single molecule PCR reveals similar patterns of non-homologous DSB repair in tobacco and arabidopsis. PLoS One 7(2): e32255.

Luijsterburg, M. S., I. de Krijger, W. W. Wiegant, R. G. Shah, G. Smeenk *et al.*, 2016 PARP1 Links CHD2-Mediated Chromatin Expansion and H3.3 Deposition to DNA Repair by Non-homologous End-Joining. Mol. Cell 61: 547–562.

McVey, M., and S. E. Lee, 2008 MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends Genet. 24: 529–538.

Nekrasov, V., B. Staskawicz, D. Weigel, J. D. G. Jones, and & S. Kamoun, 2013 Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. Nat. Biotechnol. 31: 691–693.

Nishimasu, H., F. A. Ran, P. D. Hsu, S. Konermann, S. I. Shehata *et al.*, 2014 Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell 156: 935–949.

Osakabe, K., Y. Osakabe, and S. Toki, 2010 Site-directed mutagenesis in Arabidopsis using custom-designed zinc finger nucleases. Proc. Natl. Acad. Sci. USA. 107: 12034–12039.

de Pater, S., L. W. Neuteboom, J. E. Pinas, P. J. J. Hooykaas, and B. J. van der Zaal, 2009 ZFN-induced mutagenesis and gene-targeting in Arabidopsis through Agrobacterium-mediated floral dip transformation. Plant Biotechnol. J. 7: 821–835.

de Pater, S., J. E. Pinas, P. J. J. Hooykaas, and B. J. van der Zaal, 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., and F. Fauser, 2013a Gene targeting in plants: 25 years later. Int. J. Dev. Biol. 57: 629–637.

Puchta, H., and F. Fauser, 2013b Synthetic nucleases for genome engineering in plants: prospects for a bright future. Plant J.78: 727–741.

Qi, Y., Y. Zhang, F. Zhang, J. a. Baller, S. C. Cleland *et al.*, 2013 Increasing frequencies of site-specific

mutagenesis and gene targeting in Arabidopsis by manipulating DNA repair pathways. Genome Res. 23: 547–554.

Roerink, S. F., R. Schendel, and M. Tijsterman, 2014 Polymerase theta-mediated end joining of replication-associated DNA breaks in C. elegans. Genome Res. 24: 954–962.

Rulten, S. L., A. E. O. Fisher, I. Robert, M. C. Zuma, M. Rouleau *et al.*, 2011 PARP-3 and APLF function together to accelerate nonhomologous end-joining. Mol. Cell 41: 33–45.

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

Shirley, B. W., S. Hanley, and H. M. Goodman, 1992 Effects of ionizing radiation on a plant genome: analysis of two Arabidopsis transparent testa mutations. Plant Cell 4: 333–347.

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

Vu, G. T. H., H. X. Cao, K. Watanabe, G. Hensel, F. R. Blattner *et al.*, 2014 Repair of site-specific DNA double-strand breaks in barley occurs via diverse pathways primarily involving the sister chromatid. Plant Cell 26: 2156–2167.

Walker, J. R., R. A. Corpina, and J. Goldberg, 2001 Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. Nature 412: 607–614.

Wang, M., W. Wu, W. Wu, B. Rosidi, L. Zhang *et al.*, 2006 PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. Nucleic Acids Res. 34: 6170–6182.