ZFN-induced mutagenesis and gene-targeting in Arabidopsis through Agrobacterium-mediated floral dip transformation

Sylvia de Pater*, Leon W. Neuteboom, Johan E. Pinas, Paul J. J. Hooykaas and Bert J. van der Zaal

Department of Molecular and Developmental Genetics, Institute of Biology Leiden, Leiden University, Sylviusweg 72, 2333 BE Leiden, The Netherlands

Summary
Zinc-finger nucleases (ZFNs) are artificial restriction enzymes, custom designed for induction of double-strand breaks (DSBs) at a specific locus. These DSBs may result in site-specific mutagenesis or homologous recombination at the repair site, depending on the DNA repair pathway that is used. These promising techniques for genome engineering were evaluated in Arabidopsis plants using Agrobacterium-mediated floral dip transformation. A T-DNA containing the target site for a ZFN pair, that was shown to be active in yeast, was integrated in the Arabidopsis genome. Subsequently, the corresponding pair of ZFN genes was stably integrated in the Arabidopsis genome and ZFN activity was determined by PCR and sequence analysis of the target site. Footprints were obtained in up to 2% of the PCR products, consisting of deletions ranging between 1 and 200 bp and insertions ranging between 1 and 14 bp. We did not observe any toxicity from expression of the ZFNs. In order to obtain ZFN-induced gene-targeting (GT), Arabidopsis plants containing the target site and expressing the ZFN pair were transformed with a T-DNA GT construct. Three GT plants were obtained from ~3000 transformants. Two of these represent heritable true GT events, as determined by PCR, Southern blot analysis and sequencing of the resulting recombined locus. The third plant showed an ectopic GT event. No GT plants were obtained in a comparable number of transformants that did not contain the ZFNs. Our results demonstrate that ZFNs enhance site-specific mutagenesis and gene-targeting of Agrobacterium T-DNA constructs delivered through floral dip transformation.

Keywords: Arabidopsis, double-strand break, floral dip transformation, gene-targeting, mutagenesis, zinc-finger nucleases.

Introduction
Genetic modification of plants is now routinely performed. Transformation can be done by various methods and vectors including Agrobacterium tumefaciens. It has been observed that transgenes integrate at fairly random positions and in variable copy numbers in the plant genome through non-homologous recombination (NHR). This may cause position effects (like silencing of transgenes) and unintended mutations of genes at the integration site. Therefore, it would be an advantage if integration could be targeted to a specific locus. Such gene-targeting (GT) would also be of great advantage for the modification or inactivation of genes in the plant genome. GT can be achieved by homologous recombination (HR). This process is efficient in yeast but very rare in most higher eukaryotes, like animals and plants. Estimates of GT frequencies in several different plant species vary from $10^{-4}$ to $10^{-6}$ (Paszkowski et al., 1988; Lee et al., 1990; Offringa et al., 1990; Halfter et al., 1992; Hrouda and Paszkowski, 1994; Miao and Lam, 1995; Risseeuw et al., 1995; Hanin et al., 2001). Gene-targeting frequency can be increased by the introduction of a targeted DNA double strand break (DSB) near the site of the desired recombination event. In plants,
intrinsically error prone. When active ZFNs are present, the this particular DNA repair pathway. Therefore, NHEJ is contrast to HR, no homologous sequences are required for repaired via non-homologous end-joining (NHEJ). In con- plants, one can exploit the fact that DSBs are mainly prints in the ZFN target sequence were detected in Arabid- recent also been instrumental for precise GT, a process where a DNA sequence inserts at the induced DSB site, preferably via a double HR event. This has by now been demonstrated in Drosophila embryos (Beumer et al., 2006), in human cells (Urnov et al., 2005; Lombardo et al., 2007; Moehle et al., 2007), in zebrafish (Doyon et al., 2007) as well as in plant protoplasts and plant cell suspensions. As for the plant experiments, ZFN-induced GT was demonstrated by precise repair of defective reporter genes by means of integration of specific DNA con- structs at the ZFN target site (Wright et al., 2005; Cai et al., 2008). Very recently, it was demonstrated that ZFN technology can be instrumental for HR-mediated targeted integration (Shukla et al., 2009) and mutation (Townsend et al., 2009) of endogenous genes in maize and tobacco. In most studies, the GT repair constructs and ZFNs expres- sion constructs were co-delivered to cell suspensions (Cai et al., 2008; Shukla et al., 2009) or protoplasts (Wright et al., 2005; Townsend et al., 2009) via a variety of direct DNA transformation procedures. For plants, it will be of great interest to further develop ZFN-induced GT by means of DNA transfer via the widely used bacterial vector Agrobacterium tumefaciens. Thus far, in respect of ZFN- induced GT, Agrobacterium has been used only for tobacco cell cultures (Cai et al., 2008).

Here we report the use of ZFNs to create DSBs for the introdution of site-specific mutations and to enhance the frequency of gene-targeting in Arabidopsis by means of the simple floral dip method (Clough and Bent, 1998). A reporter-based assay was developed exploiting PZF domains that had been shown to bind to chromatin-embedded DNA. The functionality of the ZFNs was first tested in a yeast-based test system. Subsequently, foot- prints in the ZFN target sequence were detected in Arabid- opsis plants expressing the ZFNs. It was shown that floral dip transformation of such plants with an incoming homologous T-DNA GT construct can indeed lead to precise GT events.

This was demonstrated for the first time through the use of the rare cutting meganuclease I-SceI, resulting in an increase in GT frequency by two orders of magnitude (Puchta et al., 1996). In order to create a DSB at a pre- determined site in the genome, zinc-finger nucleases (ZFNs) are rapidly emerging as the tools of choice (Porteus and Carroll, 2005; Camenisch et al., 2008). The current generation of ZFNs combines the nonspecific cleavage domain of the FokI restriction enzyme and a specific DNA binding domain with several C2H2 zinc-fingers (ZFs) to pro- vide cleavage specificity. Efficient cleavage of the target site requires dimerization of the FokI cleavage domain (Bitinaite et al., 1998; Smith et al., 2000; Mani et al., 2005). Therefore, two ZFN subunits are typically designed to recognize the target sequence in a tail-to-tail configuration and the DSB is then introduced within a spacer sequence which is located between the binding sites of the two polydactyl zinc-finger (PZF) domains.

Each individual ZF present within a PZF consists of a stretch of ~30 amino acids, stabilized by a zinc ion that binds a particular three-base DNA sequence (triplet). At the present date, a series of ZF modules has been created for recognition of most of the 64 possible triplets (Sega et al., 1999; Dreier et al., 2001, 2005; Liu et al., 2002). Although not all interactions are robust enough for this purpose, the availability of the ZF lexicon and the recently published ZFN selection method OPEN (Maeder et al., 2008) in principle allows the construction of effective ZFNs. A minimum of three ZFs are required per PZF domain, but more complex PZF domains recognizing longer target sites usually possess increased DNA binding specificity as well as higher affinity for the cognate recognition site. In previous work, we developed a convenient repetitive cloning method that facilitates the construction of extended PZF domains. Different types of PZF designs were critically evaluated for their ability to bind to chromosomal chromatin-embedded DNA sequences in the yeast Saccharomyces cerevisiae (Neuteboom et al., 2006). Subsequently, it was also demonstrated that PZF domains constructed via this method were highly effective in Arabidopsis, enabling novel mutant screens (Lindhout et al., 2006) and in vivo labelling of chromosomal target sites (Lindhout et al., 2007).

For site directed mutagenesis (SDM) in animals and plants, one can exploit the fact that DSBs are mainly repaired via non-homologous end-joining (NHEJ). In con- trast to HR, no homologous sequences are required for this particular DNA repair pathway. Therefore, NHEJ is intrinsically error prone. When active ZFNs are present, the cycle of cutting and repairing the ZFN target site continues until an imperfect NHEJ-mediated repair event results in a mutation or footprint within the target site, which pre- vents recognition or subsequent cleavage by the ZFNs. Such a NHEJ-based mutagenesis strategy was developed in Drosophila (Bibikova et al., 2002) and was also shown to be an efficient mutagenesis method in Arabidopsis, tobacco and maize (Lloyd et al., 2005; Maeder et al., 2008; Shukla et al., 2009; Tovkach et al., 2009).

In addition to SDM via the generation of site-specific DSBs and their subsequent imperfect repair, ZFNs have been demonstrated for precise GT, a process where a DNA sequence inserts at the induced DSB site, preferably via a double HR event. This has by now been demonstrated in Drosophila embryos (Beumer et al., 2006), in human cells (Urnov et al., 2005; Lombardo et al., 2007; Moehle et al., 2007), in zebrafish (Doyon et al., 2007) as well as in plant protoplasts and plant cell suspensions. As for the plant experiments, ZFN-induced GT was demonstrated by precise repair of defective reporter genes by means of integration of specific DNA con- structs at the ZFN target site (Wright et al., 2005; Cai et al., 2008). Very recently, it was demonstrated that ZFN technology can be instrumental for HR-mediated targeted integration (Shukla et al., 2009) and mutation (Townsend et al., 2009) of endogenous genes in maize and tobacco. In most studies, the GT repair constructs and ZFNs expres- sion constructs were co-delivered to cell suspensions (Cai et al., 2008; Shukla et al., 2009) or protoplasts (Wright et al., 2005; Townsend et al., 2009) via a variety of direct DNA transformation procedures. For plants, it will be of great interest to further develop ZFN-induced GT by means of DNA transfer via the widely used bacterial vector Agrobacterium tumefaciens. Thus far, in respect of ZFN- induced GT, Agrobacterium has been used only for tobacco cell cultures (Cai et al., 2008).
Results

ZFN functional assay in yeast

For the production of DSBs, we made use of the PTF and E2C PZF DNA binding domains that were previously shown to bind to their target site in chromatin-embedded DNA (Neuteboom et al., 2006). These PZFs were coupled to the FokI nuclease domain with a 21 amino acid linker, having about the same length as a linker used previously (Smith et al., 1999) (Figure 1a). To determine the functionality of PTFFOK and E2CFOK ZFNs for DSB formation in chromatin-embedded DNA, these ZFNs were expressed in yeast and an artificial chromosomal target site, with a 6 or 16 bp spacer between the PZF recognition sequences, were analysed for the presence of footprints. These spacers lengths were chosen since Smith et al. (2000) demonstrated that 6 bp to 35 bp spacers were cleaved by ZFNs possessing the long amino acid linker in between the PZF and the FokI cleavage domain mentioned above. When also our ZFNs would efficiently digest longer target sites in addition to sites with a 6 bp spacer, it was reasoned that the use of a 16 bp spacer would reveal this. An EcoRI restriction site was included in the spacer sequences (Figure 1b), which may be lost upon imperfect repair of the ZFN-induced break. The 16 bp spacer also contained a PvuII and a SpeI restriction site. Expression of the ZFNs was confirmed by Western blotting using anti-FLAG antibodies (results not shown). Only when both PTFFOK and E2CFOK were present and when 6 bp separated the recognition sites, PCR products were found that were lacking the EcoRI site. The sequences of cloned EcoRI resistant PCR products exhibited deletions ranging from 1 to 26 bp (Figure 1c). These results showed that the ZFNs constructed for our study are functional for DSBs formation in chromatin-embedded yeast DNA, recognizing a target site with 6 bp spacing rather than 16 bp.

ZFN-induced mutagenesis in Arabidopsis plants

For ZFN-induced mutagenesis in Arabidopsis, an assay was developed utilizing PTFFOK and E2CFOK and their target sequences. First, a T-DNA containing the ZFNs target sequences with 6 bp spacing was introduced in Arabidopsis (Figure 2a). One plant line was selected containing one full-length copy of the T-DNA. By means of TAIL-PCR, the integration site of the T-DNA was found to be located on chromosome 4 between At4g22750 and At4g22760 (Figure 2b). Both flanking regions were analysed by PCR and sequencing. Constructs encoding PTFFOK and E2CFOK both under control of the Rps5a promoter, which is primarily active in meristematic cells and early embryos (Weijers et al., 2001), the constitutive 35S promoter, or the tamoxifen-inducible 35S promoter (pINTAM construct; Friml et al., 2004) were stably integrated in the homozygous target plant line. The presence of both ZFN genes was checked by PCR using primers amplifying the coding region of both ZFNs and subsequent HindIII digestion of the DNA fragment (Figure 3). Most plant lines contained both ZFNs as determined by the presence of five DNA fragments. A selection of these plant lines was analysed by RT-QPCR for expression of the ZFNs. Relative expression levels were determined with primers in both the E2C and PTF PZF domains, and showed similar expression levels of both ZFNs (results not shown). To compare expression levels of ZFN constructs with different promoters, primers in the FokI nuclease domain were used (Figure 4). As expected, the relative expression levels were low in lines.

Figure 1 Outline of the yeast ZFN test system. A. Linker sequences present between the PZF domains and the FokI nuclease domain of the ZFNs used here (1) compared with the linker used by Smith et al., 1999 (2). B. The ZFNs target locus was inserted in the PDC6 locus, containing the target sequences (boxed) for the PTF and E2C PZF domains in a tail-to-tail orientation with 6 or 16 bp spacing containing an EcoRI site (underlined). C. Footprints in the yeast target locus. Deleted nucleotides are indicated by dots. Nucleotides which can be placed left or right of the deletion are shown in grey.
that contained the Rps5a constructs (4.4 ± 2.5). The constitutive 35S promoter constructs gave up to 200-fold higher expression (727 ± 365), which was comparable with its tamoxifen-induced version (874 ± 1020).

Subsequently, the activity of the ZFNs was evaluated. Although footprints could be detected in primary transformed plants, we routinely used T2 seedlings as a source for DNA isolation since these could be grown, induced (in case of the pINTAM-ZFNs) and analysed at any convenient time point. Genomic DNA isolated from a mixture of ten seedlings was pre-digested with EcoRI and used as template in PCR reactions (Figure 5). PCR products derived from all different types of transgenic plants, harboring ZFNs driven by the different promoters, were digested with EcoRI and EcoRI resistant fragments were cloned and sequenced (Figure 6a). The target plant line that did not

Figure 2 Outline of the Arabidopsis GT system. (a) The T-DNA containing the target locus containing the P \(_{\text{Rps5a}}\)-gfp/gus reporter gene including the ZFN target sequence and the P \(_{\text{nos}}\)-ppt selection marker was stably integrated. ZFNs were stably integrated in this line. Their functionality was tested by detection of footprints in their target sequence. GT was induced by introduction of a GT T-DNA construct with homology to the target locus (P \(_{\text{nos}}\)P \(_{\text{Rps5a}}\) and GUS), but missing the coding region of the ppt selection marker and having the gfp coding sequence replaced by hpt. GT will result in the recombined locus as shown. HindIII sites generating the DNA fragment that was detected after GT by the hpt probe on the Southern blots (Figure 7b) are shown. (b) Sequence of the chromosome 4 plant DNA flanking the T-DNA containing the target locus in the selected Arabidopsis line as determined by TAIL-PCR and PCR. LB, left border; RB, right border.
contain ZFNs did not reveal EcoRI resistant PCR products (Figure 5). A quantitative analysis of footprint frequency in ZFN expressing plant lines was impossible due to the cycle of EcoRI digestion and PCR. It was observed, however, that footprints were harder to detect when the relatively weak Rps5A promoter was driving ZFN expression. Most EcoRI resistant DNA fragments contained deletions ranging from 1 to 80 bp. Several larger deletions, up to about 200 bp, were found as well as small insertions (1–14 bp), the latter mostly accompanied by deletions. Often the repair had taken place in a region containing microho-
Figure 6  Sequences of EcoRI resistant target sites. (a) EcoRI digested DNA from several plant lines with different promoter-ZFN constructs was used for PCR and EcoRI resistant products were cloned and sequenced. ZFN binding regions are boxed and the EcoRI site is underlined. Deleted nucleotides are indicated by dots and inserted nucleotides in bold. Nucleotides which can be placed left or right of the deletion are shown in grey. The sizes of the deletions and insertion are indicated. (b) Sequences of EcoRI resistant PCR products obtained without pre-digesting the genomic DNA with EcoRI.
mology (Figure 6, grey boxes), suggesting that in these cases microhomology mediated end-joining (MMEJ; McVey and Lee, 2008) had been used for repair.

In order to get more reliable insight in the frequency of mutations within the target sequence, PCR products amplified using undigested DNA templates from a mixture of T2 seedlings expressing 35S-ZFNs (plant line 3) were cloned and sequenced directly, thus without any selection by means of EcoRI digestion. Out of 240 cloned PCR fragments, 5 were missing the EcoRI site. The sequences of

Figure 6 (Continued)
these five clones showed deletions of 1–48 bp (Figure 6b). This means that in about 2% of the cells the target sequence contained footprints, which were likely formed after ZFN-induced DSB formation followed by incorrect repair via NHEJ. The majority of the target sites was unaltered and was therefore still potential substrates for ZFN cleavage. Remarkably, no aberrant phenotypes were observed in any of the plant lines expressing the ZFNs, indicating the absence of ZFN toxicity.

ZFN-induced gene-targeting in Arabidopsis plants

Since the ZFNs were shown to create DSBs in their target sequence, the next step was to perform gene-targeting experiments at this locus. Due to the fact that the majority of potential ZFN target sites was still intact in cells of T2 seedlings, even when the ZFN pair was under control of the constitutive 35S promoter, we decided to investigate whether it was possible to achieve GT by introducing a homology carrying construct with the commonly used floral dip procedure. With a potentially active ZFN pair expressed in flower buds and undamaged target sites still present in most of the cells, it should be sufficient to employ an Agrobacterium strain carrying no more than a single T-DNA for which HR-mediated integration at the target locus can be demonstrated. The experimental design of the GT assay is shown in Figure 2a. A GT T-DNA is introduced by floral dip transformation into a plant line already containing the target locus as well as the genes encoding the ZFNs, the latter located on one T-DNA with the kanamycin selection marker. We chose to use plants with ZFN expression driven by the Rps5A promoter in order to avoid unwanted constitutive activity of the ZFN in differentiated tissues. The fact that this promoter is predominantly active in meristematic cells, such as present during flower and gamete formation as well as early embryos (Weijers et al., 2001), ensures that ZFN expression will be at hand during the process of floral dip transformation. The GT T-DNA was designed to be similar to the sequence at the target locus, but the gfp coding region was replaced by the hpt coding region and the ppt coding region, present in the target locus, was omitted. Upon integration via HR the hpt gene will replace the gfp coding region and will be linked to ppt and to the chromosome 4 plant DNA flanking the target locus. Transgenic seeds from the floral dip transformation were selected on hygromycin. It should be noted that random integration of the GT construct carrying the Rps5A-driven hpt gene also results in hygromycin resistant seedlings. Therefore, PCR was used to detect GT events. Pools of 20 hygR plants were analysed with primers in hpt and chromosome 4. Of the 152 pools (3040 plants) that were analysed, three pools contained one plant each that produced a PCR product of the correct size with primers in the hpt gene and the downstream flanking region of chromosome 4 (Figure 7a), indicating that the hpt coding region was integrated via homologous recombination. In two of the three plants (GT1 and GT3) the correct band was also detected with primers in the upstream flanking region of chromosome 4 and the hpt coding region. These two plants represent so-called true gene targeting events (TGT). In GT2, homologous recombination in the region upstream of the hpt gene linking the ppt gene and hpt gene, could not be detected, indicating that integration at this end did not occur via HR. Longer extension times also did not result in PCR products with ppt and hpt specific primers. Therefore, it seems that integration of the T-DNA at this end also did not occur via NHEJ. This plant represents an ectopic gene targeting (ETG) event. In all three plants the original target locus was still present, indicating that the plants were heterozygous. No GT events were detected in 2860 hygR plants that did not contain the ZFNs.

To confirm the GT events, Southern blot analysis was performed. Genomic DNA of the plants GT1, GT2, and GT3 was digested with HindIII and probed with hpt (Figure 7b), detecting GT T-DNA fragments with one recognition site in the T-DNA and one recognition site in the flanking plant DNA (Figure 2a). A 4.6 kb hybridizing fragment expected for a GT event was detected in DNA from all three plants. In the lanes with DNA from both GT1 and GT3 many additional hybridizing bands were visible. These bands probably represent randomly integrated T-DNAs. The additional, small band (1 kb) in the lanes with GT2 DNA probably represents a partial T-DNA.

Progeny of the GT plants was analysed by PCR for the segregation of the gfp and hpt genes (Figure 7c). In GT2, segregation was only observed for the hpt locus. Selection on hygromycin of GT2 progeny showed that GT2 contained one active hpt gene (180 hygR : 53 hygS; 3 : 1). All hygR progeny plants still contained the gfp gene, indicating that HR had taken place between the target locus and the GT T-DNA through one-end invasion and gene conversion. Apparently, this was followed by release of the T-DNA, completed with the newly acquired sequence of chromosome 4, and integration elsewhere in the genome. However, segregation of gfp and hpt and thus of the target locus and the recombined GT locus was observed in GT1 and GT3. Three classes of GT1 and GT3 progeny
plants were obtained; homozygous gfp, homozygous hpt and heterozygous plants. Southern blot analysis of the progeny of GT1 and GT3 showed that some of the additional bands representing randomly integrated T-DNAs were not present anymore (Figure 7b). In one GT1 progeny plant all of the extra bands had been lost (Figure 7b; GT1 plant 19). This means that the extra copies were not linked to the GT locus and that they could be lost by selfing. PCR analysis of the ZFN genes showed that in some GT1 progeny plants these genes were lost through segregation (results not shown). One of these is the heterozygous GT1 plant 19, which also did not contain randomly integrated T-DNAs. By selfing this plant we obtained homozygous plants in which both target loci had been replaced by the GT T-DNA and in which the randomly integrated T-DNAs as well as the ZFNs were not present anymore (results not shown). In all hygR GT2 progeny plants, the band of expected size (4.6 kb) is present as well as the small band (Figure 7b), which is probably linked to the ectopic GT locus.

Finally, the full-length T-DNA inserts comprising the target locus and the recombined GT locus in GT1 and GT3 were amplified with primers in chromosome 4 and cloned. Since the plants that were used were heterozygous for the target locus and the recombined GT locus, two different T-DNAs were cloned. Recombined GT clones from

Figure 7 GT analysis in Arabidopsis. (a) PCR on primary GT plants with primers in chromosome 4 (SP295 or SP289) flanking the target T-DNA combined with primers in the hpt gene (SP284 or SP292) or gfp gene (SP251 or SP293) to detect homologous recombination. (b) Southern blot analysis of DNA from primary transformants (P) and progeny GT plants (numbers). Genomic DNA was digested with HindIII and blots were probed with the hpt gene. Molecular weight markers (M) are Lambda (HindIII) (top panel) or Lambda (EcoRI/HindIII). The lanes in panel GT1 are derived from several different blots and band intensities cannot be compared. (c) PCR on progeny of GT plants using a primer in chromosome 4 flanking the target locus (SP289) combined with primers in the hpt gene (SP292) or gfp gene (SP293) to detect segregation. Plants homozygous for the GT locus are underlined. For GT2 23 hygR progeny plants and one hygS progeny plant (16) were used for PCR.
both GT1 and GT3 were sequenced from the N-terminal part of the \textit{prot} coding region to the N-terminal of the \textit{hpt} coding region and from the C-terminal part of the \textit{hpt} coding region to chromosome 4. The sequences were identical to those expected from precise homologous recombination between the target locus and the GT T-DNA (results not shown). Clones of the original target locus from both the GT1 and GT3 lines were sequenced and both still contained intact ZFN recognition sequences, indicating that either this chromosome had not been broken by the ZFNs or that repair had occurred here via precise NHEJ. Taken together, these results show that GT occurred in plants with ZFN activity, whereas no GT was detected in a comparable number of plants without ZFNs.

### Discussion

The ZFNs that were generated at the onset of this study contained 6-fingered (6ZF) PZF domains. Hence, in theory, any dimeric ZFN that can be composed within a cell expressing one or both of these ZFN should preferably interact with a 2 × 18 bp DNA sequence, given that the spacing in between the 18 bp sequence allows the dimer to do so. Any particular need for such long ZFN recognition sites has thus far not been demonstrated. In fact, most other studies targeted 2 × 9 or 2 × 12 bp half-sites, corresponding with two PZFs each consisting of 3 or 4 ZFs (Lloyd \textit{et al.}, 2005; Urnov \textit{et al.}, 2005; Wright \textit{et al.}, 2005; Beumer \textit{et al.}, 2006; Doyon \textit{et al.}, 2007; Lombardo \textit{et al.}, 2007; Moehle \textit{et al.}, 2007; Cai \textit{et al.}, 2008; Maeder \textit{et al.}, 2008; Tovkach \textit{et al.}, 2009; Townsend \textit{et al.}, 2009).

Thus far only one study made use of ZFNs with longer DNA binding domains containing 5 or 6 ZF (Shukla \textit{et al.}, 2009). Our choice was inspired by our earlier findings that 6ZF PZF s were superior regarding their in vivo interaction with chromosomal DNA in yeast cells (Neuteboom \textit{et al.}, 2006). Moreover, in the same study the binding characteristics of 6ZF PZF domains PTF and E2C were carefully analysed. For the purpose of investigating ZFN-induced mutagenesis and GT in Arabidopsis, the performance of the \textit{PTFFOK} and \textit{E2CFOK} ZFN pair was also first evaluated in a yeast assay to ensure that they were active on chromatin-embedded chromosomal DNA.

The linker between the PZF and the \textit{FokI} nuclease domain was 21 amino acids, similar to the length of the linker used by Smith \textit{et al.} (1999). The typical footprints resulting from imperfect DSB repair within and around a unique 2 × 18 bp target site were only detected with the 6 bp spacer between the ZFN recognition half sites and not with the 16 bp spacer. Recently, it was reported that ZFNs containing linkers of up to 20 amino acids bind also to target sequences with longer spacers (Händel \textit{et al.}, 2009). We did not observe this, indicating that probably the amino acid sequence of the linker also determines the affinity for the target site.

Once a single target locus was introduced into Arabidopsis and homozygous lines were obtained, the strategy that was used to finally obtain GT in those plants encompassed a two-step procedure: after raising plants harbouring ZFNs, these very plants were subsequently transformed with a GT donor construct. So far, ZFN-induced GT in eukaryotic cells has almost exclusively been accomplished by the simultaneous introduction of DNA molecules encoding ZFNs and an incoming donor GT construct with homology to the site of interest where the DSB should occur. As far as plant cells are concerned, the target cells used for transformation aimed at ZFN-induced GT thus far were protoplasts (Wright \textit{et al.}, 2005; Townsend \textit{et al.}, 2009) or cultured cells (Cai \textit{et al.}, 2008; Shukla \textit{et al.}, 2009). Although this does not at all affect any of the conclusions regarding successful ZFN-mediated GT events, the very processes of protoplasts isolation or tissue culture steps needed to support growth or regenerations of cells and tissues might have a detrimental effect on the normal state of DNA metabolism, the state as it is maintained in unstressed cells or tissues. These tissue culture approaches may lead to somaclonal variations and chromosomal rearrangements (Kaeppler \textit{et al.}, 2000; Mohan Jain, 2001; An \textit{et al.}, 2005). Floral dip transformation, as it is routinely performed with intact Arabidopsis plants, does not necessitate any kind of regeneration of deliberately prepared cells or tissues. This method in fact prevents the tissue culture steps that may induce unwanted changes in the genome. Although the method is mainly used for Arabidopsis, efforts have resulted in protocols for floral dip transformation of other species, like wheat (Zale \textit{et al.}, 2009). Furthermore, the female gametes that are transformed via floral dip transformation are in a developmental phase just before or after meiosis (ovule primordia, megasporocytes or megagametocytes) (Desfeux \textit{et al.}, 2000) and might therefore be equipped with highly active HR DNA repair systems, more so than somatic cells or cells experiencing stress by artificial manipulations.

As mentioned above, the expression of ZFNs can potentially be accompanied by toxic side-effects, possibly caused by high expression or off-target binding of the ZFNs (Cornu \textit{et al.}, 2008). However, it cannot be excluded that in cases where detrimental effects are observed, these are
at least partly the consequence of in vitro manipulations of target cells rather than a harmful effect of ZFNs. In this respect, our data that footprints were observed in primary transformants expressing ZFNs as well as in seedlings derived from these plants without noticeable negative effects upon plant growth and development that could be attributed to ZFN expression are very much encouraging regarding the feasibility of using ZFNs in plants. To our knowledge, these are the first data providing evidence that multicellular organisms can very well tolerate expression of active ZFNs. For all ZFN expressing plant lines maintained in our laboratory, the offspring is of perfectly normal wild type appearance.

It can be argued that the absence of any detrimental effects of ZFN expression also has to be attributed to the fact that only a small percentage (2%) of the cells from plants expressing the ZFNs contains footprints, even when the ZFNs were expressed using the 35S promoter. This may imply that the ZFNs were unstable or not very active. Although ZFN protein was easily detectable in yeast, we could not detect ZFN protein in plants on Western blots using anti-FLAG antibodies. In addition, full-length ZFN mRNA levels were very low in plants. This could mean that plant specific RNases or proteases degrade our ZFN mRNAs or proteins. Recently, it was shown that by modulating ZFN protein levels, ZFN toxicity can be reduced (Pruett-Miller et al., 2009). Thus, low protein levels may also be a reason for the absence of toxic effects. Another possibility, which might very much be the consequence of using unstressed plant material as also mentioned above, is that most DSBs that were formed were rapidly repaired without leaving any footprint due to the fact that the repair system remains very active during the normal processes of embryo-, seedling- and plant development. In this respect, it is of particular interest to note that even in mammalian cell extracts, the majority of DSBs are repaired without leaving any footprint (Kuhfittig-Kulle et al., 2007). The footprints that were detected in our study were all rather typical mutations as they are created via NHEJ or MMEJ repair of DSBs, thus resulting in deletions and insertions. The size of the deletions observed was limited. However, these results greatly depend on the size of the PCR fragments that were isolated and cloned. It cannot be excluded that also larger deletions had been created upon DNA repair.

With healthy plant lines expressing ZFNs, but with the large majority of target sites still unaltered, it was possible to perform another round of floral dip transformation aimed at GT by means of a HR event at a newly generated DSB. For the GT experiments, plants with Rps5a-dri-

ven ZFNs, with by estimation much less than 2% of footprints within the ZFN target sites, were used for floral dip. As already mentioned above, the Rps5A promoter is rather active in dividing cells, so also in developing flowers, gametes, zygotes and early embryos (Weijers et al., 2001). Therefore, during floral dip of these plants a significant number of the cells that are transformed with the GT construct may have a ZFN-induced DSB.

We observed a GT frequency of $10^{-3}$ in our GT assay in the presence of functional ZFNs. Considering the estimates of GT frequency in several different plant species ranging from $10^{-4}$ to $10^{-6}$ (Paszkowski et al., 1988; Lee et al., 1990; Offringa et al., 1990; Halfter et al., 1992; Hrouda and Paszkowski, 1994; Miao and Lam, 1995; Risseeuw et al., 1995; Hanin et al., 2001), our data suggest a 10- to 1000-fold increase in GT frequency due to ZFN expression. Since we did not find any GT event in plants without ZFNs, it is currently not possible to know the exact increase of GT frequency by ZFNs in our test system. In literature, much higher ZFN-induced GT frequencies were reported (Wright et al., 2005; Cai et al., 2008; Shukla et al., 2009; Townsend et al., 2009). However, it should be realized that a comparison cannot be made as the methods that were used differ very much from ours, involving different plant species, in vitro cultured plant material, different methods of DNA delivery and different genomic loci. In future experiments, direct selection of GT events will enable us to screen larger numbers of transformants, with and without ZFNs. This should allow us to determine with precision the fold increase in GT by ZFN-induced DSBs by means of the floral dip method. The results presented here corroborate recent findings in very different systems that ZFN-induced mutagenesis and ZFN-induced GT are very promising technologies for making precise changes in the plant genome at any given locus. Of particular importance in this respect is the finding that even prolonged exposure of the cells of multicellular organisms to ZFNs with sufficient specificity seems to present no particular fundamental problem, not even for their offspring.

Experimental procedures

Construction of yeast strains with the zinc-finger nuclease target site and vectors for the expression of ZFNs

The plasmid pINT1 (Meijer et al., 1999) was digested with BamHI and annealed primers for target sites with a 6 bp spacer (GATCTC-
ATCGCGCCGCCCACATCGAATTCGGGCCGGAGCCCGACATG and GATCCACTGCGGCTCAGGGCCCGGAATTCAGATGTCTCCGGC- CGCGCATG) or a 16 bp spacer (GATCTCATCGGGCCGCGCC- GATCGATGGAATTCGCGGGTGGCGCCGGACGCGCAGTG and GATCCACTGCGGCTCAGGGCCCGGAATTCAGATGTCTCCGGC- CGCGCATG) were ligated and confirmed by sequencing. Plasmids containing the binding site in the correct orientation were linearized with SacI and NcoI, isolated from agarose gel and transformed to the yeast strain 21R.Jmel1 (Melcher et al., 2000) by a PEG/LiAc transformation method (Gietz et al., 1992). Yeast colonies were checked by PCR for correct insertion at the PDC6 locus.

Plasmid pET-15b:ΔQNK-FnI (Smith et al., 1999) (kind gift of Dr Chandrasegaran, John Hopkins University, Baltimore, MD, USA) was digested with BamHI and Spel and the ~0.6 kb band containing the coding region of the FokI endonuclease domain was cloned into BglII and Spel-digested pSKN-SgrAI, pSKN-SgrAI-PTF1 and pSKN-SgrAI-E2C1 (Neuteboom et al., 2006). These PZF will be further referred to as PTF and E2C, respectively. *Escherichia coli* DH5α strains containing the pSKN-SgrAI plasmids with a FokI domain were always grown in the presence of 20 mM glucose to suppress protein expression.

For expression in yeast, the different ZFN constructs were cloned as NotI fragments into similarly digested p425ADH or p426ADH vectors. p426ADH, which has been described previously (Neuteboom et al., 2006), allows selection on medium lacking uracil. The vector p425ADHl, which allows selection on medium containing the coding region of the *P. pastoris* PDC6 locus.

### Yeast nuclease assay

p425ADH and p426ADH plasmids, either without insert or containing ZFN fusions, were transformed separately as well as simultaneously to yeast strains containing the artificial target site and colonies were selected on minimal medium supplemented with the appropriate amino acids. For the nuclease assay, yeast cells were grown in selective liquid medium to an O.D.600nm of 0.8–1.0. Cells from 1 mL of culture were washed twice with 1 mL TE buffer. Subsequently, samples were heated at 94 °C for 5 min with occasional shaking. Cell debris was removed by centrifugation for 5 min and 2 μL of the denatured genomic DNA sample was used for PCR amplification. PCR was performed with REDTaq™ polymerase (Sigma-Aldrich, St. Louis, MO, USA) in 25 μL using 15 pmol each of TargetFW primer and TargetREV primer (Table 1). The presence of the EcoRI site was evaluated by restriction analysis.

### Construction of the ZFN target site, the GT repair construct, ZFN expression vectors and plant transformation

The ZFN target sequence was inserted in the *Bam*HI site of pGPTV-BAR containing the Rps5a-gfp-gus reporter (Weijers et al., 2001) resulting in pSDM3832 (target T-DNA). The gfp coding sequence (*Bam*HI/Scal fragment) was replaced by hpt (*BglII/*BamHI fragment).

### Table 1 PCR primers

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>targetFW</td>
<td>ATTCGTTTGGATACACTACTAATGCC</td>
<td>ZFN target sequence (yeast) sense</td>
</tr>
<tr>
<td>targetREV</td>
<td>AACATTATATCGAAAAAACCCTTG</td>
<td>ZFN target sequence (yeast) antisense</td>
</tr>
<tr>
<td>SP258</td>
<td>CGACTACAAGGACGACGAGC</td>
<td>N-terminus ZFN sense</td>
</tr>
<tr>
<td>SP259</td>
<td>CCTCTAAAGGTATATGCGCC</td>
<td>C-terminus ZFN antisense</td>
</tr>
<tr>
<td>SP272</td>
<td>CCTTAATGAATGTTGGTAAAG</td>
<td>Fok domain sense</td>
</tr>
<tr>
<td>SP273</td>
<td>GTTTGTGAGCGGGACAGTATTAC</td>
<td>PTF antisense</td>
</tr>
<tr>
<td>SP274</td>
<td>GTCAGAGAAAAAGGACCCTTCC</td>
<td>E2C antisense</td>
</tr>
<tr>
<td>SP275</td>
<td>CTCGCCGTCTCTCCTCATC</td>
<td>ZFN target sequence (Arabidopsis) sense</td>
</tr>
<tr>
<td>SP278</td>
<td>CTTAGAAGAGCTGTCGTGCTCT</td>
<td>ZFN target sequence (Arabidopsis) antisense</td>
</tr>
<tr>
<td>SP283</td>
<td>GTAGTGTTGACAGATGGTC</td>
<td>ptt sense</td>
</tr>
<tr>
<td>SP284</td>
<td>CAGCAGATTCCTCCGGCCTCC</td>
<td>hpt antisense</td>
</tr>
<tr>
<td>SP285</td>
<td>CTTAGAATGGTGCCGAGGTTGTG</td>
<td>Chr4 right flank sense</td>
</tr>
<tr>
<td>SP292</td>
<td>CCGAGAGCCAAAGAAATAGAG</td>
<td>hpt sense</td>
</tr>
<tr>
<td>SP293</td>
<td>GTGCCGCCCTGAGCACAAAGACC</td>
<td>gfp sense</td>
</tr>
<tr>
<td>SP295</td>
<td>GACAGAGGATATACACATACGAGTTTGTGG</td>
<td>Chr4 left flank antisense</td>
</tr>
<tr>
<td>ROC3.3</td>
<td>CCACAGGCTCCGCTGCTTCC</td>
<td>ROC1 sense</td>
</tr>
<tr>
<td>ROC5.2</td>
<td>GAAGCGAGACGCGGGTGGATGTC</td>
<td>ROC1 antisense</td>
</tr>
<tr>
<td>MC141</td>
<td>CGATTCCGGAAGTGGATGTC</td>
<td>Hpt sense</td>
</tr>
<tr>
<td>MC142</td>
<td>GTGCCGCCATCATCCTATTCC</td>
<td>Hpt antisense</td>
</tr>
<tr>
<td>XB1208FW</td>
<td>AATTCGATATCCCTCATGATG</td>
<td>Linker pCambia1200del</td>
</tr>
<tr>
<td>XB1208RV</td>
<td>GTGGAATATGAAATT</td>
<td>Linker pCambia1200del</td>
</tr>
<tr>
<td>NOS 1</td>
<td>GATTGAATACCTGTGGCCGGTCTTT</td>
<td>Tail PCR</td>
</tr>
<tr>
<td>NOS 2</td>
<td>GCATGACGTATTTATGAGATGG</td>
<td>Tail PCR</td>
</tr>
<tr>
<td>NOS 3</td>
<td>CGCAAACATGATAAAATTACCAGG</td>
<td>Tail PCR</td>
</tr>
<tr>
<td>AD3</td>
<td>(A-)/GTGNAG(A-)/JANCANAGA</td>
<td>Tail PCR</td>
</tr>
</tbody>
</table>
from pGPTV-HPT (Becker et al., 1992), resulting in pSDM3833 (positive control). The GT construct was created by cloning the NheI/Pmel fragment of pSDM3833 in pCambia1200del (XbaI/EcoRV), creating pSDM3834. pCambia1200del is a derivative of pCambia1200 with the Xmrnl-BstXI fragment containing 35S-hpt replaced for a linker (Xb 1200FW/RV) containing an EcoRI site. PTFFOK and E2CFOK were cloned as NotI fragments in pGPTV derivatives (Becker et al., 1992; kan or hpt selection markers), in which the ZFNs are driven by the RpsSa promoter (pSDM3835, pSDM3836), the 35S promoter (pSDM3838, pSDM3839) or the tamoxifen-inducible 35S promoter (pSDM3840, pSDM3941) (Friml et al., 2004). An Xba linker was cloned in the Pmel site in pSDM3836 and the RpsSa-PTFFOK gene from pSDM3836 was cloned as an XbaI fragment in pSDM3835 containing RpsSa-E2CFOK, resulting in pSDM3837. The two ZFN genes in pSDM3837 were placed in inverted orientation. Plasmids are listed in Table 2.

Plant vectors were introduced in Agrobacterium tumefaciens AGL1 (Lazo et al., 1991) by electroporation.

Arabidopsis plants (ecotype Col-0) were transformed via the floral dip method (Clough and Bent, 1998) and primary transformants were selected on MA solid medium without sucrose ral dip method (Clough and Bent, 1998) and primary transformants were placed in inverted orientation. Plasmids are listed in Table 2.

DNA isolation and PCR analysis

Seedlings, leaves or flowers were disrupted to a powder under liquid N2 in a TissueLyser (Retch, Haan, Germany). The powder was mixed with 500 mL CTAB extraction buffer (2% CTAB (N- cetyl-N,N-trimethyl ammonium bromide), 1.4 NaCl, 20 EDTA, 100 Tris–HCl, pH 8.0) supplemented with RNase (0.2 mg/mL), timentin (100 µg/mL) and the appropriated antibiotics (ppt 15 µg/mL; hpt 15 µg/mL; km 30 µg/mL) for selection of transformed plants.

<table>
<thead>
<tr>
<th>Table 2 Plasmids</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmid number</td>
</tr>
<tr>
<td>pSDM3832</td>
</tr>
<tr>
<td>pSDM3833</td>
</tr>
<tr>
<td>pSDM3834</td>
</tr>
<tr>
<td>pSDM3835</td>
</tr>
<tr>
<td>pSDM3836</td>
</tr>
<tr>
<td>pSDM3837</td>
</tr>
<tr>
<td>pSDM3838</td>
</tr>
<tr>
<td>pSDM3839</td>
</tr>
<tr>
<td>pSDM3840</td>
</tr>
<tr>
<td>pSDM3841</td>
</tr>
</tbody>
</table>

RT-PCR

For expression analysis of ZFN constructs, pools of ten days old seedlings were frozen in liquid N2. For expression of pINTAM-ZFNs, seedlings were incubated for 48 h in liquid cotMS alone or cotMS supplemented with 1 µM tamoxifen for induction of the promoter. Tissue was disrupted to a powder under liquid N2 in a TissueLyser (Retch, Haan, Germany). RNA was isolated using RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA). Residual DNA was removed with DNA-free kit (Ambion, Austin, TX, USA). cDNA was produced on 0.5 µg RNA using the iScript cDNA synthesis kit (Bio-Rad, Hercules, CA, USA) and PCR was performed on 0.25 µL cDNA using the iScript SYBR Green Supermix (Bio-Rad, Hercules, CA, USA) in a final volume of 25 µL with 0.4 µm of each primer using a DNA Engine Thermal Cycler (MJ Research, Ramsay, MI, USA) equipped with a Chromo4 real-time PCR detection system (Bio-Rad, Hercules, CA, USA). Expression of ZFNs was analysed with primers SP258 and SP274 for the E2C ZFN domain, with primers SP258 and SP273 for the PTF PFZ domain and with primers SP272 and SP259 for the FoXI domain. Normalization of relative gene expression was based on expression of the housekeeping gene roc1 (primers ROC3.3, ROC5.2).

Nuclease assay in Arabidopsis

Genomic DNA was digested with EcoRI and target sites were amplified with Phusion polymerase (Finnzymes, Espoo, Finland) using primers SP250 and SP251. PCR products were digested with EcoRI and analysed by gel electrophoresis. EcoRI resistant fragments were isolated from 5% polyacrylamide gels (in some cases amplified by PCR to obtain enough material) and cloned in pCt1.2 (Fermentas, Burlington, Ontario, Canada). DNA sequences were determined by ServiceXS (Leiden, the Netherlands).

Southern blot analysis

Plant DNA (5 µg) was digested with HindIII and separated in 0.7% agarose gels, blotted onto Hybond-N and hybridized in DIG easy hyb (Roche Diagnostics, Mannheim, Germany) supplemented with 50 µg/mL herring sperm DNA with a hpt probe, labelled in a PCR reaction with primers MC141 and MC142 using DIG-labelling mix (Roche Diagnostics, Mannheim, Germany). After 16–20 h, blots were washed twice with 2x SSC; 0.1% SDS at room temperature and three times with 0.2xSSC; 0.1% SDS at 65 °C. Detection was performed using the DIG wash and block buffer set and CDP-star (Roche Diagnostics, Mannheim, Germany) according the manufacturers protocol.
Acknowledgements

We want to thank Dr S. Chandrasegaran for the plasmid containing the FokI coding sequence and Martin Brittin for preparation of the figures. This work was financially supported by a grant from the Dutch ministry of economical affairs (GenTap IS054064) and a grant from the EU (Recbreed KBBE-2008-227190).

References


© 2009 The Authors
Journal compilation © 2009 Blackwell Publishing Ltd, Plant Biotechnology Journal, 7, 821–835


