Plant Mol Biol (2009) 69:699–709
DOI 10.1007/s11103-008-9449-7

Targeted transgene integration in plant cells using designed zinc finger nucleases

Charles Q. Cai · Yannick Doyon · W. Michael Ainley · Jeffrey C. Miller · Russell C. DeKelver · Erica A. Moehle · Jeremy M. Rock · Ya-Li Lee · Robbi Garrison · Lisa Schulenberg · Ryan Blue · Andrew Worden · Lisa Baker · Farhoud Faraji · Lei Zhang · Michael C. Holmes · Edward J. Rebar · Trevor N. Collingwood · Beth Rubin-Wilson · Philip D. Gregory · Fyodor D. Urnov · Joseph F. Petolino

Received: 22 October 2008 / Accepted: 14 December 2008 / Published online: 27 December 2008

Abstract Targeted transgene integration in plants remains a significant technical challenge for both basic and applied research. Here it is reported that designed zinc finger nucleases (ZFNs) can drive site-directed DNA integration into transgenic and native gene loci. A dimer of designed 4-finger ZFNs enabled intra-chromosomal reconstitution of a disabled *gfp* reporter gene and site-specific transgene integration into chromosomal reporter loci following co-transformation of tobacco cell cultures with a donor construct comprised of sequences necessary to complement a non-functional *pat* herbicide resistance gene. In addition, a yeast-based assay was used to identify ZFNs capable of cleaving a native endochitinase gene. *Agrobacterium* delivery of a Ti plasmid harboring both the ZFNs and a donor DNA construct comprising a *pat* herbicide resistance gene cassette flanked by short stretches of homology to the endochitinase locus yielded up to 10% targeted, homology-directed transgene integration precisely into the ZFN cleavage site. Given that ZFNs can be designed to recognize a wide range of target sequences, these data point toward a novel approach for targeted gene addition, replacement and trait stacking in plants.

C. Q. Cai · W. M. Ainley · R. Garrison · L. Schulenberg · R. Blue · A. Worden · L. Baker · B. Rubin-Wilson · J. F. Petolino (✉)
Dow AgroSciences, LLC, 9330 Zionsville Road, Indianapolis, IN 46268, USA
e-mail: jfpetolino@dow.com

Y. Doyon · J. C. Miller · R. C. DeKelver · E. A. Moehle · J. M. Rock · Y.-L. Lee · F. Faraji · L. Zhang · M. C. Holmes · E. J. Rebar · T. N. Collingwood · P. D. Gregory · F. D. Urnov
Sangamo BioSciences, 501 Canal Blvd., Suite A 100, Richmond, CA 94804, USA

Keywords Double strand DNA breaks · Homology-directed repair · Site-directed transgene integration · Zinc finger nucleases

Introduction

The ability to introduce foreign DNA into a predetermined location within the plant genome would provide a powerful tool for basic studies of plant gene function and enhance the development of new crop varieties (Kumar and Fladung 2001; Puchta 2002; Reiss 2003; Hanin and Paszkowski 2003). Although well established in yeast (Aylon and Kupiec 2004) and moss (Trouiller et al. 2007), targeted transgene integration remains a significant challenge in higher plants (Iida and Terada 2005; Endo et al. 2007). Site-specific transgene integration occurs at a very low frequency in plant cells as compared to random integration, even when the introduced DNA contains large stretches of sequence with homology to host DNA (Paszkowski et al. 1988; Offringa et al. 1990; Lee et al. 1990; Miao and Lam 1995; Kempin et al. 1997; Hanin et al. 2001). Attempts to enhance targeted integration efficiencies in plants have included the use of negative selection markers (Wang et al. 2001; Terada et al. 2002, 2007) and plants genetically engineered to exhibit modified recombination processes (Shalev et al. 1999; Reiss et al. 2000). These efforts notwithstanding, site-specific transgene integration in plant cells remains a challenge (Puchta 2005).

In this regard, substantial increases in the frequency of targeted integration in a broad range of animal and plant model systems have been observed following the induction of a DNA double strand break (DSB) at a specific genomic location in host cells, which stimulates a native cellular process, homology-directed DSB repair (Porteus and Carroll

2005). Naturally occurring endonucleases whose recognition sites are rare in the plant genome (Jasin 1996) have been used in this manner to drive transgene integration into a target sequence previously transferred into the plant genome via random integration (Puchta et al. 1993; Puchta 1998; Siebert and Puchta 2002; D′Halluin et al. 2008). These studies highlight the potential of targeted DSB induction to stimulate site-specific transgene integration in plant cells, though the challenge of introducing a DSB in an endogenous locus remains. The solution to targeted induction of a DSB at a native genomic location is provided by zinc finger nucleases (ZFNs). The C_2H_2 zinc finger protein domain has been used as a scaffold for sequence-specific DNA binding (Pavletich and Pabo 1991) and ZFNs produced by fusing zinc finger protein domains with a sequence-independent nuclease domain derived from the Type IIS restriction endonuclease *Fok*I (Kim et al. 1996). Engineered ZFNs have been used to drive high-efficiency targeting to an endogenous genomic locus in transformed (Moehle et al. 2007) and primary human cells (Lombardo et al. 2007).

Initial attempts at using ZFNs in plants have been promising (Lloyd et al. 2005; Wright et al. 2005; Maeder et al. 2008). A construct carrying a ZFN gene under the control of an inducible promoter along with its corresponding recognition sequence was stably integrated into *Arabidopsis* and shown to introduce targeted mutations resulting from non-homologous end joining at the recognition site at frequencies averaging 7.9% among induced progeny seedlings (Lloyd et al. 2005). Similarly, among 66 tobacco plants regenerated from protoplasts transformed with a ZFN designed to cleave at the SuRA locus, three displayed single base pair deletions at the target site resulting from non-homologous end joining repair (Maeder et al. 2008). Tobacco cells, containing a pre-integrated, non-functional reporter gene missing 600 bp directly flanking a zinc finger recognition sequence, when co-transformed with constructs containing a corresponding ZFN gene and donor DNA homologous to the pre-integrated sequence comprising the missing 600 bp, showed evidence of homology-directed repair of the reporter gene (Wright et al. 2005).

In the present study, an engineered tobacco cell culture system with a pre-integrated reporter construct—containing dual partial, non-functional reporter genes flanked by multiple ZFN binding sites separated by large stretches of non-homologous sequence—was used to show that ZFNs can facilitate site-specific cleavage and transgene integration following co-transformation with an appropriately designed donor DNA. It was shown that ZFNs can drive the excision of large segments of DNA from an integrated target construct concomitant with ZFN-mediated insertion of a heterologous DNA stretch specified by the donor DNA. Importantly, these findings were extended to the targeting of an endogenous gene encoding an endochitinase protein (Fukuda et al. 1991). A yeast-based assay system (Doyon et al. 2008) was used to identify ZFN pairs for the cleavage of tobacco endochitinase. ZFN designs identified by screening in this system, in combination with donor DNA constructs containing homologous sequences and an herbicide resistance selectable marker, drove targeted integration into the endogenous endochitinase gene locus.

Materials and methods

Constructs

Target reporter construct

A hygromycin phosphotransferase (*hpt*) gene (Waldron et al. 1985) was used as a selectable marker for the initial transformation. Two Green Fluorescent Protein (*gfp*) gene fragments (Evrogen Joint Stock Company, Moscow, Russia) flanked a *β*-glucuronidase (*uidA*) expression cassette. The *N. tabacum* RB7 matrix attachment region (MAR) (Thompson et al. 1997) and the *A. thaliana* 4-coumaryl synthase (4-CoAS) intron-1 (Locus At3g21320, GenBank NC 003074) served as homologous sequences-1 and -2, respectively, and a 3′ fragment of the *S. viridochromogenes* phosphinothricin phosphotransferase (*pat*) (Wohlleben et al. 1988) gene was included for in vitro selection in subsequent transformations (Fig. 1a). Zinc finger binding sites comprising recognition sequences consisting of inverted repeats to which zinc finger–FokI fusion proteins can bind as homodimers (Fig. 1b) were integrated into the target construct. Each binding site contained four tandem repeats of the recognition sequence of the particular zinc finger–FokI fusion protein so that each binding site was ∼200 bp in size to ensure that the recognition sequences were accessible to the zinc finger–FokI fusion protein in the complex chromatin environment. One binding site (ZFN-T1, Fig. 1a) was fused with the *uidA* coding sequence at the N-terminus. Two copies of a second binding site (ZFN-T2, Fig. 1a) flanked the 5′ and 3′ *gfp* gene fragments. The 5′ and 3′ *gfp* gene fragments overlapped by 540 bp providing homology within the target sequence and a stop codon was inserted at the 3′ end of the 5′ *gfp* fragment to ensure no functional *gfp* transcription from the target sequence.

Zinc finger nuclease constructs

pZFP-1 and pZFP-2 were assembled from an archive of in vitro-selected modules (Isalan and Choo 2001; Urnov et al. 2005) as previously described (Doyon et al. 2008) and yielded the following ZFP moieties: pZFP-1—DRSTLIE SSSNLSR RSDDLSK DNSNRIK; pZFP-2—F1 RSDDLSK DNSNRIK RSDALSV DNANRTK. Zinc finger proteins were designed,

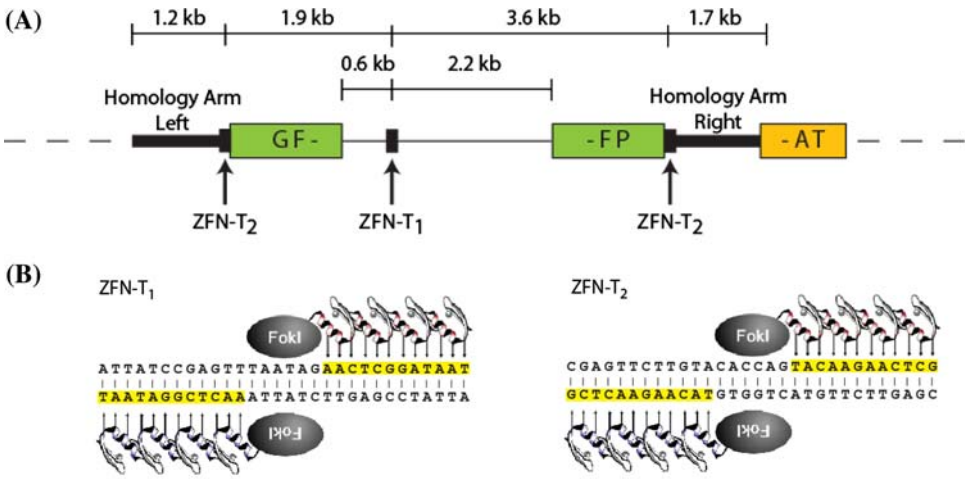

Fig. 1 Schematic of the target reporter and zinc finger recognition sequences. **a** Target reporter construct with tandem, overlapping, partial *gfp* gene fragments, a 3′ *pat* selectable marker gene fragment, left and right homologous sequences and zinc finger binding sites. **b** Recognition sequences for ZFN-1 and ZFN-2

assembled, and screened for cleavage activity against the coding sequence of the endochitinase gene CHN50 (Fukuda et al. 1991) as previously described (Doyon et al. 2008) to yield the following ZFP moieties: ZFP-CHN50-L, TCCGA CCAGGAG, RSANLAR RSDNLRE DRSNLSR DSS DRKK; ZFP-CHN50-R, TCGGACGAGGCC, DNRDLIR RSDDLSR DRSNLSR RNDDRKK. Assembled ZFPs were cloned in-frame as NH_2-terminal fusions to the catalytic domain of FokI (Kim et al. 1996). The zinc finger–FokI fusion protein genes were driven by a CsVMV promoter and 5′ UTR (Verdaguer et al. 1996) (Fig. 3a, b). A dual expression cassette (pZFN-CHN50) containing ZFN-CHN50-L and ZFN-CHN50-R flanking a 2A sequence (Fang et al. 2005) was used for endogenous endochitinase gene cleavage (Fig. 3c).

Donor constructs

(i) The reporter donor DNA construct (pDONOR) consisted of homologous sequence-1, a full-length *A. thaliana* ubi10 promoter (Callis et al. 1990), 299 bp of 5′ 'partial' *pat* gene coding sequence (Wohlleben et al. 1988) that complements the 3′ partial *pat* gene in the target sequence and homologous sequence-2. The homologous sequences in pDONOR were identical to those in pTARGET. A construct, pSELMARK, containing an intact *pat* gene expression cassette consistent with the expected homologous recombinants included the *A. thaliana* 4-CoAS intron-1 (Locus At3g21320, GenBank NC 003074) inserted at the 299/300 bp of a complete *pat* coding sequence (Wohlleben et al. 1988). (ii) The endogenous endochitinase gene donor construct (pCHITINASE) contained a 1,504 bp fragment of the CHN50 locus (Fukuda et al. 1991) generated by PCR amplification from tobacco BY2 genomic DNA using primers: 5′-CAATGTGGTTCGCAGGCGGG-3′ and 5′-GCTCATTAACACATCTATTGTGGACAAAGTC-3′. A novel StuI site was introduced between bases 750 and 751 of this fragment using the QuikChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) so that a *pat* gene cassette could be inserted into this site and cloned into pZFN-CH50 to generate pCHITINASE.

Transformation

Target reporter sequence integration

Target sequences were stably integrated into BY2 tobacco cell suspension cultures via *Agrobacterium* transformation. BY2 cultures (obtained from Jun Ueki of Japan Tobacco, Iwata, Shizuoka, Japan) were maintained in media containing LS basal salts (PhytoTechnology Labs, Shawnee Mission, KS, #L689), 170 mg/l KH_2PO_4, 30 g/l sucrose, 0.2 mg/l 2,4-D and 0.6 mg/l thiamine-HCl at a pH of 6.0. The BY2 cells were sub-cultured every 7 days by adding 0.25 ml PCV to 50 ml of LS-based medium maintained in 250-ml flasks on a rotary shaker at 25°C and 125 rpm. In order to generate transgenic BY2 cell cultures with integrated target sequences, a flask of a four-day post subculture suspension was divided into 10–12 four ml aliquots which were co-cultivated in 100 × 25 mm Petri dishes with 100 µl *Agrobacterium* strain LBA4404 harboring pTARGET grown overnight to an OD_{600} ~1.5. Dishes were wrapped with Nescofilm® (Azwell Inc., Osaka, Japan) and incubated at 25°C without shaking for 3 days after which excess liquid was removed and replaced with 11 ml of LS medium containing 500 mg/l carbenicillin. Following resuspension of the tobacco cells, 1 ml suspension was dispensed onto 100 × 25 mm plates of LS medium containing 500 mg/l carbenicillin and 200 mg/l hygromycin solidified with 8 g/l TC agar (PhytoTechnology, Shawnee Mission, KS), and incubated unwrapped at 28°C in the dark. This resulted in 120–144 selection plates for a single treatment. Individual hygromycin-resistant isolates appeared 10–14 days after plating and were transferred to individual 60 × 20 mm plates (one isolate per plate) where they were

maintained under selection as callus on a 14-day sub-culture schedule until needed for analysis and subsequent re-transformation experiments.

Integration into target reporter locus

Transgenic cell cultures containing a single, full-length integrated copy of the target sequence were selected and used to re-initiate suspension cultures by placing ~250–500 mg of callus tissue into 40–50 ml of LS medium containing 100 mg/l hygromycin and sub-culturing every 7 days as described above. *Agrobacterium*-mediated transformation of the target cell cultures were performed as described above. Following co-cultivation, the cells were plated out on LS medium containing 500 mg/l carbenicillin and 15 mg/l Bialaphos®. Individual Bialaphos®-resistant isolates appeared 2–4 weeks after plating and were transferred to individual 60 × 20 mm plates (one isolate per plate) where they were maintained under selection as callus on a 14-day sub-culture schedule until needed for analysis.

Targeted integration into the endogenous CHN50 locus

Agrobacterium strain LBA4404 harboring pCHITINASE was used to transform tobacco BY2 suspension culture cells and Petite Havana leaf discs. BY2 suspension culture cells were maintained, transformed and bulked up as previously described. Leaf discs (1 cm^2) cut from Petite Havana tobacco plants aseptically grown on MS medium (Phytotechnology Labs, Shawnee Mission, KS, #M524) and 30 g/l sucrose in PhytaTrays (Sigma, St. Louis, MO) were floated on an overnight culture of *Agrobacterium* (OD_{600} ~1.5), blotted dry on sterile filter paper and then placed onto the same medium with the addition of 1 mg/l indoleacetic acid, 1 mg/l benzyamino purine. Following 48 h of co-cultivation, leaf discs were transferred to the same medium with 5 mg/l Basta™ and 250 mg/l cephotaxime. After 3 weeks, individual plantlets were transferred to MS medium with 10 mg/l Basta™ and 250 mg/l cephotaxime for an additional 2 weeks prior to leaf harvest.

Analysis

Reporter gene expression

Transgenic tobacco callus tissue and suspension cultures initiated from the selected events transformed with pTARGET were analyzed for *uidA* expression by incubating 50 mg samples in 150 µl of assay buffer for 24–48 h at 37°C (Jefferson et al. 1987). Functional *gfp* expression following intra-chromosomal recombination was monitored for 7 days following transformation using a fluorescence stereomicroscope (Leica MZ FLIII, Wetzlar, Germany) with a 470 ± 40 nm excitation filter and a 525 ± 50 nm barrier filter.

PCR analysis of the integrated target and recombinant events

Genomic DNA was extracted from tobacco callus, cell suspension cultures and leaf tissue using DNeasy 96 Plant kit (Qiagen, Valencia, CA USA) and quantified using PicoGreen ds DNA Quantitation kit (Molecular Probes, Eugene, Oregon USA). An aliquot of 2 µl of extracted genomic DNA was checked through agarose gel electrophoresis to ensure the DNA quality. To confirm integration of the full length target sequence in the selected events, nested PCR was performed using primer pair P1 (AGAGTG TGAGATACATGAATTGTCGGGC) and P2 (TGGACAG ACCCGTTCTTACACCGGACT) in the initial PCR, and P3 (TTCCGAGTCTGTAGCAGAAGAGTGAGG) and P4 (TTGGGATGGGATTGAGCTT AAAGCCGG) in the nested PCR using Takara LA Taq Polymerase (Takara, Japan). To further confirm the sequence in the selected target lines, multiple nested PCR reactions were performed. In the initial PCR reaction, the full length target sequence was amplified with Takara LA Taq Polymerase (Takara, Japan) using primer pair P1/P2. The initial PCR reaction was then diluted 100-fold and used as a DNA template in the subsequent three nested PCR reactions. These three reactions covered the entire target sequence using primer pairs P5 (GCTCTTGGACTTGTGAATTGTTCCGCC) and P6 (CGAAGGGAGCATAATAGTTACATGC), P7 (ACCAG AAAATGCTGAAAACCCGGC) and P8 (CGTATCCACG CCGTATTCGGTGATGATAAT), P9 (ACAAGCCGAAA GAACTGTACAGCG) and P10 (AGACCATGCTCAAG GTAGGCAATGTCC), respectively, with a few hundred base pair sequence overlap of each other. The nested PCR reactions were carried out using PfuTurbo® DNA Polymerase (Stratagene, La Jolla, CA). The three PCR fragments were then cloned into pCR-BluntII TOPO vector (Invitrogen, Carlsbad, CA) and sequenced by Cogenics (Houston, TX). To confirm the reconstituted *gfp*, fluorescing tissues were selectively sub-cultured and analyzed by PCR amplification of the *gfp* coding sequence using primer pair P11 (ATGCCCGCCATGAAGATCGAGTG) and P12 (AAGG CGTGCTGGTACTCCACGAT) using Takara EX Taq polymerase (Takara, Japan). All herbicide resistant, putative recombinant events were analyzed by PCR initially with primer pair of P13 (TAAGGATCCAACCATGGC TTCTCC) and P14 (AGATCTGGGTAACTGGCCTA ACTG). To further enhance the sensitivity of this assay, a nested PCR was then performed using primer pair of P15 (TACCCTTGGTTGGTTGCTGAGGTT) and P16 (GAAG GCCTATAACAGCAACCACAG). To confirm donor

DNA recombination with the target sequence as opposed to ectopic recombination, all events that were positive based on the initial PCR reaction with P13/P14 were further analyzed through nested PCR using primer pair of P13 and P17 (CAAGGTCAATACTACCAGATCTAAGA) in the first round of PCR and P14 and P15 in the nested PCR. The 5′ end of the recombined sequence was analyzed through another set of nested PCR reactions using primer pair P18 (GTGTAGAAGTACTCGCCGATAGTG) and P19 (GAC GACTAGGTCACGAGAAAGCTA) in the first PCR, and P20 (CCATGTTGGCAAAGGCAACCAAAC) and P21 (TGATAAACACGACTCGTGTGTGCC) in the nested PCR reaction. A subset of recombinant events were further analyzed by PCR using primer pair P19 and P22 (GATCT GGGTAACTGGCCTAACTGG), which amplifies a DNA fragment across the entire recombined region, from the 3′ end of the *hpt* gene to the 3′ end of the *pat* gene in the target sequence. To further confirm the sequence of the PCR products from recombinant events, the PCR fragment was purified from the gel using the QIAquick gel extraction kit (Qiagen, Valencia, CA) and then either sequenced directly using the Dye Terminator Cycle Sequencing Kit (Beckman Coulter, Fullerton, CA) or sub-cloned into the pCR2.1 TOPO vector (Invitrogen, Carlsbad, CA) then sequenced using the Dye Terminator Cycle Sequencing Kit.

Southern blot analysis

Generally, for each sample 10 μg of genomic DNA was digested with the appropriate restriction enzymes, separated on a 0.8% agarose gel, transferred and cross-linked onto a nylon membrane. Specific probes were labeled with P^{32} dCTP using the Prime-It RmT Random Primer Labeling Kit (Stratagene, La Jolla, CA) and hybridized to the membrane overnight at 65°C. The hybridized membrane was then washed in graduated increased stringency with the final wash in 0.1× SSC with 01% SDS at 65°C for 30 min. The membranes were analyzed by a BAS 1500 imaging plate scanner (Fuji Photo Film, Tokyo, Japan) or exposed on an X-rays film for a few days at −80°C. To determine the copy number of target sequence, genomic DNA was digested with *Nsi*I and sequentially probed with *pat* and *hpt* gene sequences. For further characterization of the target lines, the genomic DNA was digested with *Xho*I, *Eco*RV, *Mfe*I/*Xho*I and probed with *uidA*, *hpt*, and *pat*, respectively. For confirmation of recombinant events, genomic DNA was digested with *Ban*II, and hybridized with ^{32}P labeled probes of the 3′ partial *pat* sequence, AtuORF 24 3′ UTR and the Ubi10 promoter, sequentially. After being processed with each probe, the membranes were stripped by incubation in Regeneration Solution 1 (300 mM NaOH, 10 mM Na_2EDTA) for 10 min at room temperature followed by incubation in Regeneration Solution 2 (50 mM $NaPO_4$, 10 mM Na_2EDTA, 1% SDS) for 30 min at 65°C, then re-probed.

Results

Intra-chromosomal reporter gene correction

To evaluate the ability of a ZFN to induce a DSB in a tobacco cell genomic context, a chromosomal reporter system carrying a *gfp* gene disabled by the insertion of a 2.8 kb stretch of heterologous DNA sequence was constructed (Fig. 1a). A tandem repeat of 540 bp in the two *gfp* gene fragments served as a substrate for intra-chromosomal repair, resulting in the restoration of the *gfp* open reading frame (Fig. 2a). Transgenic suspension cultures containing a single, intact copy of the target sequence were generated and confirmed using PCR and Southern analysis (data not shown). Following re-transformation of the target-containing suspension cultures with a designed ZFN expression construct, pZFN-1 (Fig. 3a), green fluorescent foci were visible after 7 days (Fig. 2b lower). On average across several independent reporter events, ∼25 fluorescent foci per cm^2 of re-plated cells transformed with the ZFN-1 expressing construct were observed (Fig. 2b lower) compared with a background of <1 fluorescent foci per cm^2 of re-plated cells transformed with the *pat* selectable marker gene only as a negative control (Fig. 2b upper). To confirm that the observed fluorescence resulted from reconstitution of a functional *gfp* gene, a pool of fluorescing tissue segments was isolated and manually enriched through several passages of selective sub-culture. Genomic DNA was isolated from these fluorescing tissues and assayed by PCR with probes anchored on either *gfp* gene fragment. As shown in Fig. 2c, samples enriched from pZFN-1-treated, fluorescing tissues, when amplified, yielded the predicted 0.6 kb PCR product, indicating that the anticipated recombination had reconstituted a functional *gfp* gene in these tissues. An additional, 4.1 kb PCR product was also observed in the enriched samples, indicating the presence of the non-recombined reporter sequence in the cell population (Fig. 2c). This was not unexpected given the method of visual selection of fluorescing tissue used to achieve *gfp*-positive cell enrichment. These results demonstrate that ZFNs can facilitate homology-driven repair of a *gfp* reporter gene in transgenic tobacco cell cultures.

ZFN-Driven targeted integration at a chromosomal reporter locus

Using the chromosomal reporter system shown in Fig. 1a, incoming DNA was targeted to a specific chromosomal loci by co-introducing ZFN expression constructs (Fig. 3a, b) along with donor DNA that shared homology with the integrated target sequence (Fig. 4a, b). The integrated sequence contained ZFN recognition sequences (Fig. 1b)

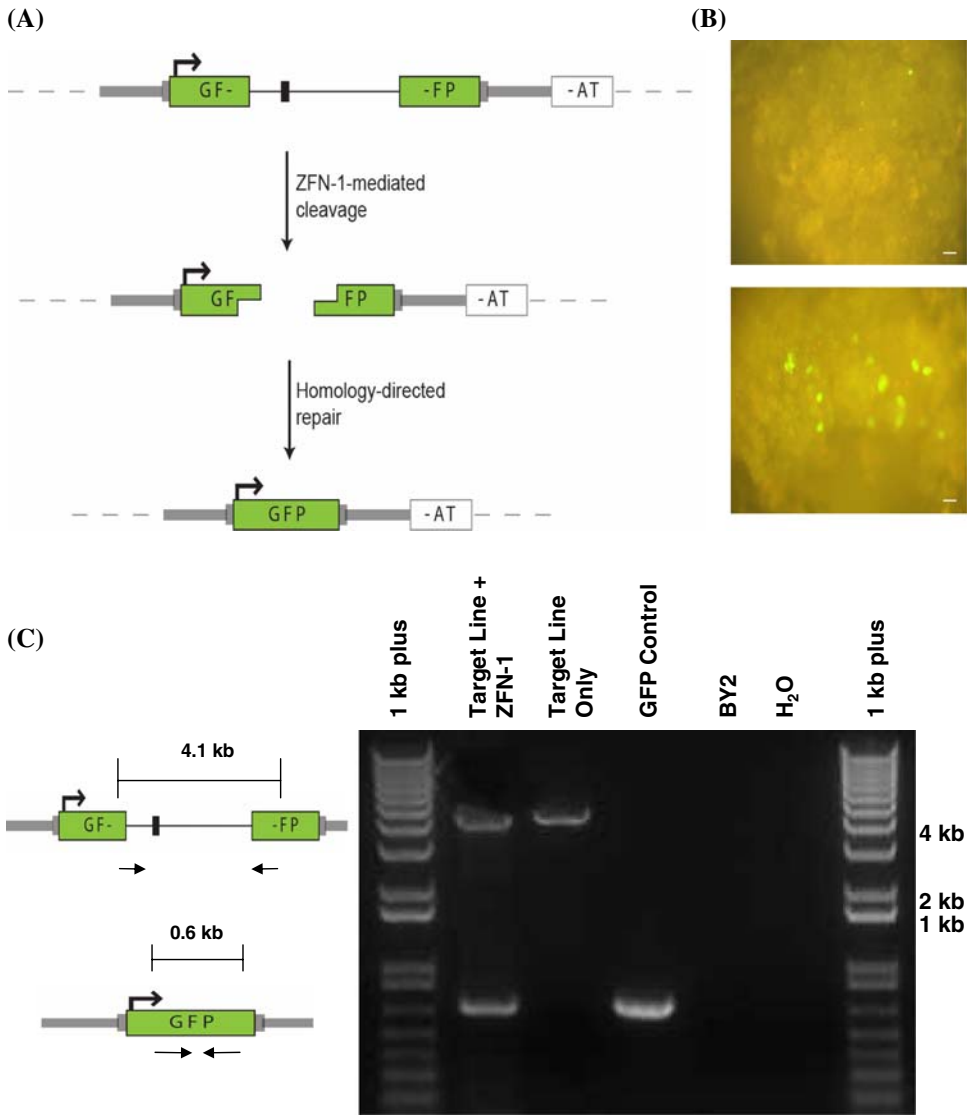

Fig. 2 Intra-chromosomal reporter gene correction. **a** Schematic of target reporter with ZFN-1 binding site and expected recombinant. **b** Fluorescent foci in BY2 suspension culture cells following *pat* (upper) and ZFN-1 (lower) gene expression and *gfp* gene reconstitution. Line = 1 mm. **c** PCR analysis of genomic DNA isolated from fluorescent tissue showing target (4.1 kb) and recombinant (0.6 kb) amplification products

which allowed for targeted cleavage at specific sites within the target. Two different integration strategies were pursued. In one approach, recognition sequences for ZFN-2 were placed directly adjacent to 1 kb stretches of sequence homologous to incoming DNA (Fig. 4a). Immediately downstream of the right homologous sequence was 256 bp representing the 3′ fragment of an herbicide resistance gene (*pat*, which confers Bialaphos®-resistance). The incoming donor DNA construct was structured such that cleavage by ZFN-2 at the target site would lead to homology-directed transfer of a 3 kb cassette that carries a promoter and the corresponding 5′ fragment of the *pat* herbicide resistance gene. A second integration strategy involved a single recognition sequence for ZFN-1 internal to a 6 kb stretch of heterologous sequence (Fig. 4b). The restoration of a functional *pat* herbicide resistance gene in this strategy would occur if the donor construct, which contains the promoter and 5′ fragment of the *pat* herbicide resistance cassette, were transferred to the chromosome via homology-directed repair using homologous sequences 3 kb distal from the ZFN cleavage site.

Use of an *Agrobacterium* strain harboring pDONOR alone, i.e. the donor DNA construct without ZFNs as a negative control, yielded no Bialaphos®-resistant isolates. Co-cultivation of transgenic suspension cultures containing a single, intact copy of the target sequence with an *Agrobacterium* strain harboring pSELMARK, a construct containing an intact *pat* selectable marker gene representing the expected recombination product, resulted in ~25 Bialaphos®-resistant isolates per ml of suspension cells). Co-cultivating target cultures with a combination of *Agrobacterium* strains harboring either pZFN-1 or pZFN-2 along with pDONOR resulted in the appearance of ~1 Bialaphos®-resistant isolate per ml of suspension cells. PCR analysis of 22 Bialaphos®-resistant isolates from a target suspension culture co-transformed with either pZFN-1

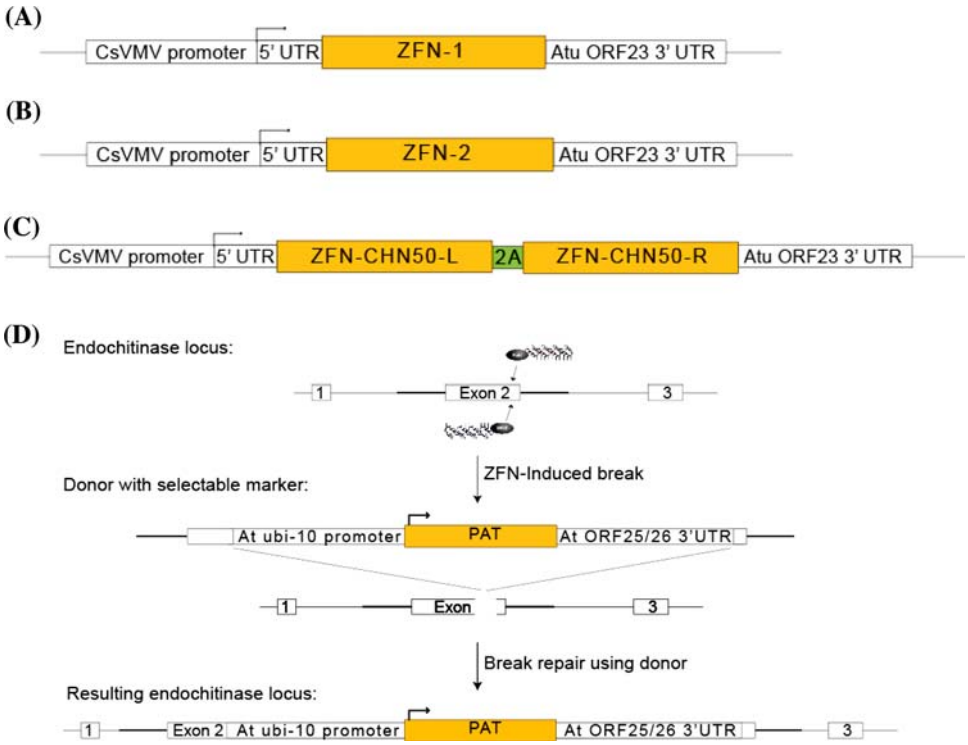

Fig. 3 Zinc finger nuclease expression cassettes. **a** pZFN-1. **b** pZFN-2. **c** pCHITINASE. **d** Endogenous endochitinase gene targeting strategy. ZFN-CHN50-L and ZFN-CHN50-R target a region in exon 2

or pZFN-2 and pDONOR is shown in Fig. 4c. The primers used anchor to the 5′ and 3′ ends of the *pat* coding sequence were designed to amplify a 2.3 kb fragment spanning the two partial *pat* gene sequences if recombination between the donor and target sequence occurred. As expected from the resistance to Bialaphos®, a 2.3 kb amplification product was observed in all 22 samples analyzed (Fig. 4c) and sequencing this DNA amplicon confirmed recombination of the 5′ and 3′ *pat* gene sequences from the donor DNA and integrated target, respectively. All events displayed sequences consistent with perfect recombination from the 5′ donor, through the homologous intron, into the 3′ target sequence. No mutations were observed in any of the events. In several of the isolates, a second prominent ~6 kb PCR amplification product was observed (Fig. 4c lanes 5, 6, 15–20). This band was shown to contain both donor and target sequences and was suspected of being the result of complex integration into the target site. Southern blot analysis following *Ban*II digestion revealed the expected 2.1 kb hybridization product from correct recombination when probed with a 3′ *pat* gene sequence (Fig. 4d). Non-recombinant target reporter constructs would yield a 3.0 kb product. A total of 18 of 22 samples displayed the ~2 kb hybridization product (lanes 1, 3–12, 14–16, 18–20 and 22). Of the four samples that did not display the ~2 kb hybridization product, two samples (lanes 17 and 21) displayed a single ~3 kb band identical to the target sequence and two samples (lanes 2 and 13) displayed a band of unexpected size (one larger and one smaller than the target sequence). Other transgenic suspension cultures containing single intact copies of the target sequence re-transformed with ZFN and donor constructs gave similar results (data not shown). These results demonstrate that co-delivery of a ZFN expression cassette and a suitably structured donor DNA construct can lead to targeted transgene integration in cultured tobacco cells.

ZFN-Driven targeted transgene integration into an endogenous locus

In order to drive transgene integration into an endogenous locus in the tobacco genome, a panel of 4-finger ZFNs, spanning the entire non-repetitive portion of an endochitinase gene, CHN50, was designed and validated using a yeast-based system. In this system, the open reading frame for a secreted form of α-galactosidase (encoded by the *MEL1* gene) was disrupted with 3 kb of sequence corresponding to the entire CHN50 genomic locus which allowed ZFNs to be screened and ranked for DNA cleavage activity. ZFNs with cleavage activity against a range of positions in the CHN50 locus were identified and a ZFN that cleaved at the extreme 3′ end of exon 2 was chosen based on this assay for use in plant cells.

A construct (pCHITINASE) containing the ZFN (Fig. 3c) along with donor DNA carrying a *pat* selectable marker driven by the CsVMV promoter flanked by 750 bp

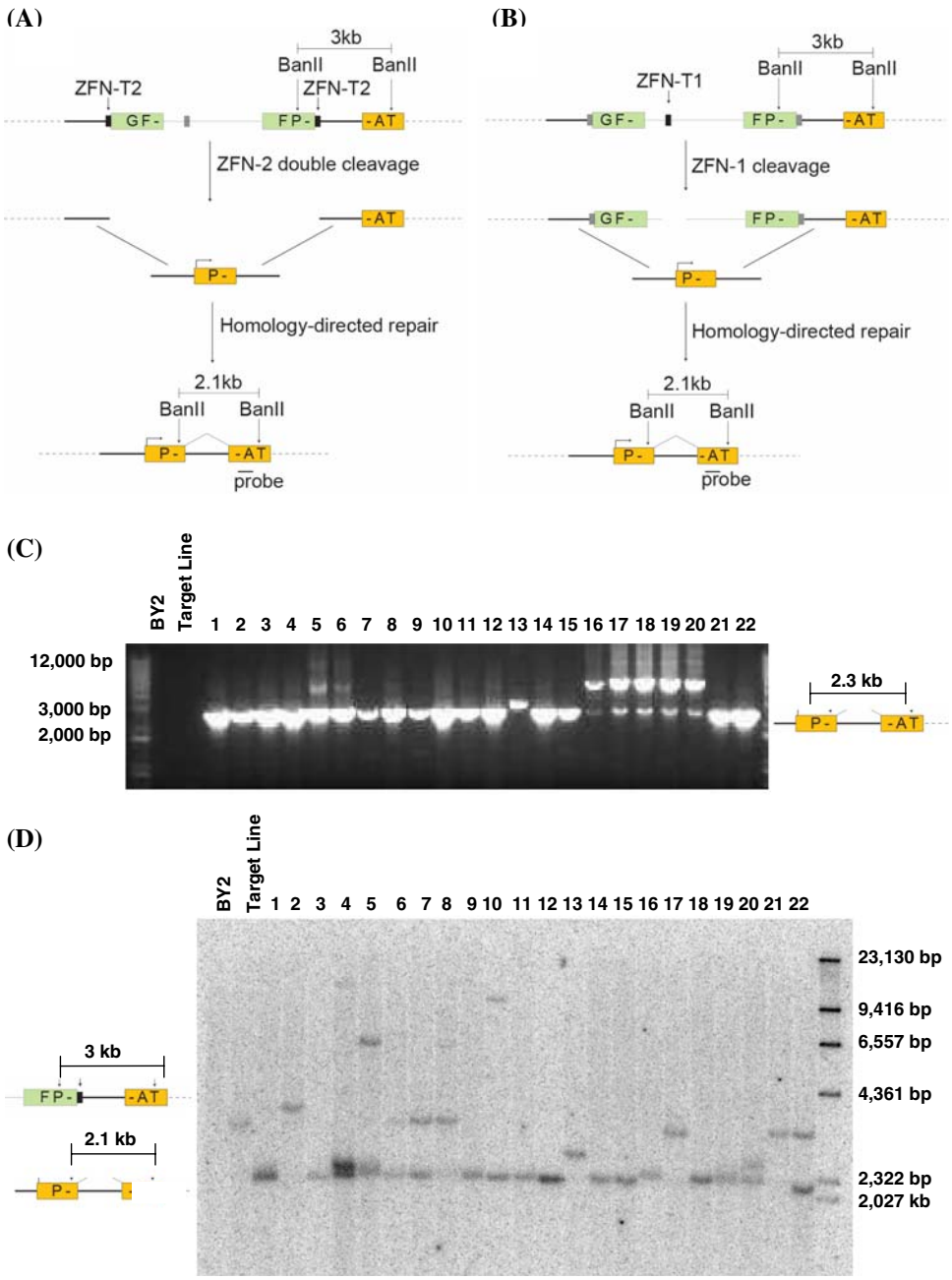

Fig. 4 Strategies for ZFN-driven targeted integration of a 3′ *pat* gene fragment into a chromosomal reporter locus and molecular analysis of targeted reporter events. **a** Targeted cleavage at two ZFN-2 binding sites. **b** Targeted cleavage at a single ZFN-1 binding site. **c** PCR analysis of genomic DNA isolated from Bialaphos®-resistant isolates showing the 2.3 kb recombinant amplification product. **d** Southern blot analysis showing targeted (2.1 kb) and non-targeted (3.0 kb) reporter loci

of sequence homologous to CHN50 (Fig. 3d) was transformed into tobacco suspension culture cells and leaf discs using *Agrobacterium*. From 231 Bialaphos®-resistant isolates obtained following co-cultivation of tobacco suspension culture cells with *Agrobacterium* harboring pCHITINASE, 12 isolates displayed both the 2.6 kb 5′ (Fig. 5a) and the 2.0 kb 3′ (Fig. 5b) amplification products expected of targeted events. Amplification products were cloned and sequenced across the donor/CHN50 gene 5′ and 3′ boundaries and displayed sequences identical to those expected to result from homology-directed repair at the target site. Similarly, from 46 transgenic plants regenerated from *Agrobacterium*-treated leaf discs, five displayed both the 5′ and 3′ amplification products indicative of targeted integration. Southern blot analysis of genomic DNA digested with the restriction enzyme *Hin*dIII (which has a single recognition sequence within the insert) using a *pat* gene probe detected a common ~5.5 kb band across the five independent targeted events and, in some cases, other unique bands resulting from random integration (Fig. 5c). Taken together, these data show that designed ZFNs can be used to drive targeted transgene integration into an investigator-specified, endogenous tobacco gene locus at a frequency up to 10% of the total events generated.

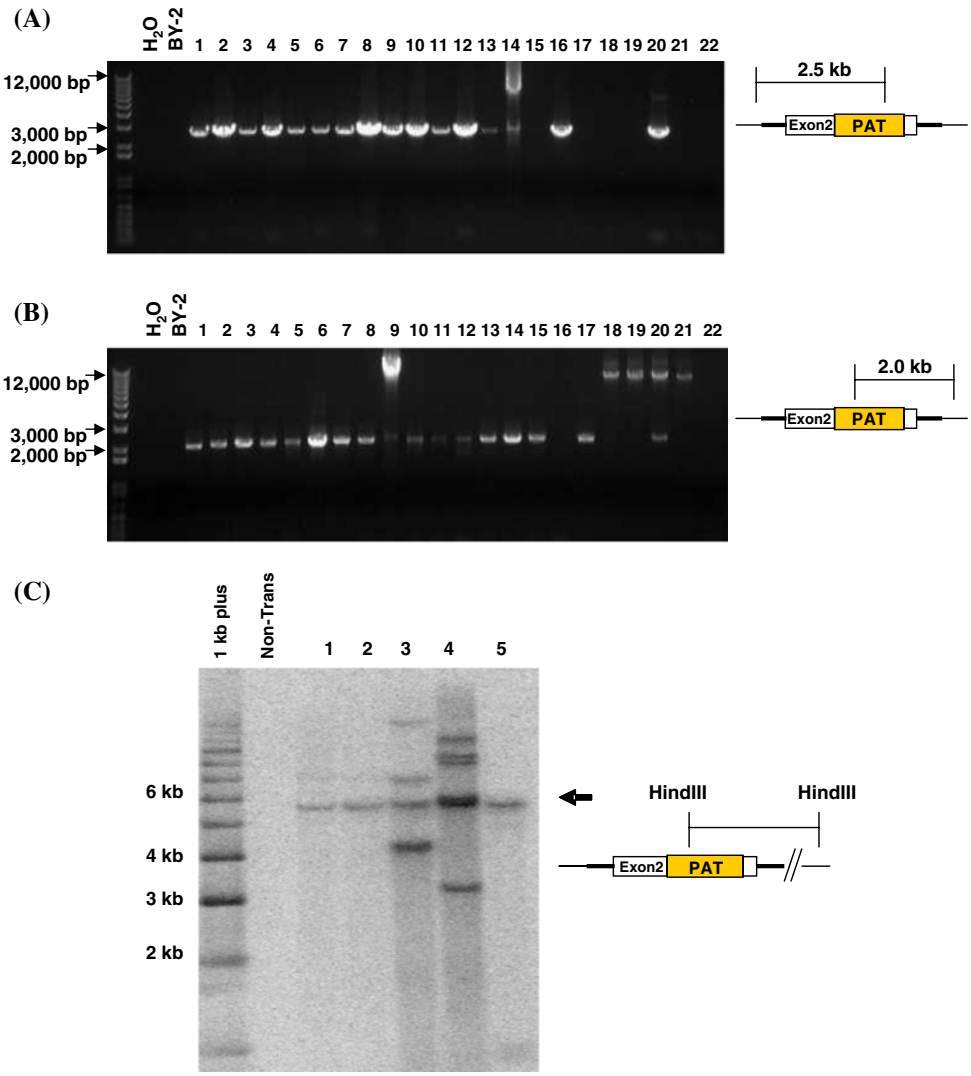

Fig. 5 ZFN-driven targeted integration into the endochitinase gene in tobacco. **a** PCR analysis of targeted lines. Amplification of the 5′ end of the insert showing a 2.5 kb amplification product. **b** PCR analysis of targeted lines. Amplification of the 3′ end of the insert showing a 2.0 kb amplification product. **c** Southern blot analysis of genomic DNA digested with *Hin*dIII showing a common ~5.5 kb band across several independent targeted events when probed with a *pat* gene sequence

Discussion

The present study describes the use of designed ZFNs to drive targeted integration of DNA into transgenic and endogenous plant genomic loci. Transgenic reporter loci stably integrated into the genome of tobacco suspension culture cells are accurately targeted with investigator-provided donor DNA following homology-directed repair. In addition, an engineered ZFN identified using a yeast screening system facilitated targeted integration of a selectable marker gene cassette into an endogenous genomic locus.

The integrated reporter system revealed several important aspects of ZFN-mediated gene targeting in plant cells. Two *gfp* gene fragments that shared 540 bp of sequence identity flanking a 4-finger designed ZFN target site provided a simple fluorescence-based assay of ZFN function. Consistent with the single strand annealing model for DNA repair (Puchta 2005), induction of a DSB at the ZFN target site followed by strand resection, annealing of complementary single strands and removal of overhanging non-complementary ends apparently resulted in deletion of ~3 kb of intervening sequence and reconstitution of a functional *gfp* gene. In addition, a 3′ partial herbicide resistance gene fragment flanked by a second ZFN binding site allowed for selection in culture following targeted DNA integration of a complementary 5′ sequence from an incoming donor construct. Approximately 6 kb of target sequence between two ZFN binding sites were excised following ZFN gene expression and replaced by 1.9 kb of sequence from a co-expressed donor DNA using 1.2 and 1.7 kb of homology directly flanking each of two induced DSBs. The resulting recombinant sequence comprised a fully functional *pat* selectable marker gene resulting in Bialaphos®-resistant colony formation. Sequencing of PCR amplified products provided evidence that repair was homology-directed at the target site. The lack of mutations at this site may have been the result of the stringent selection for functional *pat* expression. Any mutation that occurred during repair that

resulted in loss of *pat* gene function would not have survived the selection process. These results are consistent with those reported by Wright et al. (2005) who used designed 3-finger ZFNs to correct a 600 base pair deletion in a selectable marker gene with a 4.9 kb donor with 1.3 and 3.0 kb homology directly flanking the ZFN-induced DSB. Similarly, Puchta (1998) used 1 kb of homologous sequence directly flanking an I-SceI restriction site to target an herbicide resistance gene cassette to an engineered transgenic locus. In the present study, however, ZFN-mediated targeted integration at the reporter locus was observed even when the homologous sequences were 3 kb distal from a single induced DSB. In addition to highlighting the dynamic nature of plant genomes, these data suggest the ability to target incoming DNA to a relatively large region surrounding the DSB which may have important implications for ZFN target site construction when considering their application to gene editing in plants. This ability to excise and replace large stretches of integrated DNA sequence has practical implications for transgene removal, replacement and/or stacking in plant genetic engineering.

Targeted transgene integration into an endogenous locus involved ZFNs designed to bind a predetermined gene sequence screened using a yeast-based system previously shown to be suitable for this purpose in reverse genetics of zebrafish (Doyon et al. 2008). This approach is particularly useful where, in contrast to the transgenic reporter system, evaluation of ZFN activity at the endogenous locus is not trivial. A previously characterized endochitinase gene, CHN50, was chosen as a template for disruption via the addition of an investigator-specified cassette (Fukuda et al. 1991). This stress-related gene is highly expressed in stationary-phase tobacco suspension culture cells such that ability to target this site for transgene integration might be of value for recombinant protein production in tobacco cell cultures. In the present study, an herbicide resistance gene driven by a constitutive promoter flanked on each side by 750 bp of CHN50 gene sequence was co-delivered with a ZFN expression cassette to tobacco cells and leaf discs via *Agrobacterium*. Although the majority of the resulting transgenic events were the result of random integration, ~5% of the isolates generated from the suspension cultured cells and ~10% of the plant events regenerated from leaf discs were confirmed to be targeted at the expected site. Sequencing across the 5′ and 3′ endogenous gene/donor DNA boundaries confirmed homology-directed *pat* gene integration at the target site. These frequencies are sufficiently high such that the use of negative selection systems (Terada et al. 2002, 2007) may not be necessary.

These data support the growing body of evidence that ZFN-mediated DSB formation can facilitate gene targeting in plant cells via homology-directed repair and point the way toward the realization of site-specific genome editing in plants. The C_2H_2 zinc finger has proven to be remarkably adaptable to the recognition of investigator-specified sequences (Klug 2005). As shown in the present work, ZFNs can be designed against an endogenous genomic locus and used in combination with an appropriately structured donor DNA for targeted transgene integration.

Acknowledgments Thanks to Sangamo's production group for zinc finger protein assembly and Aaron Klug and Vipula Shukla for helpful comments on the manuscript.

References

Aylon Y, Kupiec M (2004) New insights into the mechanism of homologous recombination in yeast. Mutat Res 566:231–248. doi:10.1016/j.mrrev.2003.10.001

Callis J, Raasch JA, Vierstra RD (1990) Ubiquitin extension proteins of *Arabidopsis thaliana*. Structure, localization, and expression of their promoters in transgenic tobacco. J Biol Chem 265: 12486–12493

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

Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, Amora R, Hocking TD, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Amacher SL (2008) Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat Biotechnol 26:702–708. doi:10.1038/nbt1409

Endo M, Osakabe K, Ono K, Handa H, Shimizu T, Toki S (2007) Molecular breeding of a novel herbicide-tolerant rice by gene targeting. Plant J 52:157–166. doi:10.1111/j.1365-313X.2007.03230.x

Fang J, Qian JJ, Yi S, Harding TC, Tu GH, VanRoey M, Jooss K (2005) Stable antibody expression at therapeutic levels using the 2A peptide. Nat Biotechnol 23:584–590. doi:10.1038/nbt1087

Fukuda Y, Ohme M, Shinshi H (1991) Gene structure and expression of a tobacco endochitinase gene in suspension-cultured tobacco cells. Plant Mol Biol 16:1–10. doi:10.1007/BF00017912

Hanin M, Paszkowski J (2003) Plant genome modification by homologous recombination. Curr Opin Plant Biol 6:157–162. doi:10.1016/S1369-5266(03)00016-5

Hanin M, Volrath S, Bogucki A, Briker M, Ward E, Paszkowski J (2001) Gene targeting in *Arabidopsis*. Plant J 28:671–677. doi:10.1046/j.1365-313x.2001.01183.x

Iida S, Terada R (2005) Modification of endogenous natural genes by gene targeting in rice and other higher plants. Plant Mol Biol 59:205–219. doi:10.1007/s11103-005-2162-x

Isalan M, Choo Y (2001) Rapid, high-throughput engineering of sequence-specific zinc finger DNA-binding proteins. Methods Enzymol 340:593–609. doi:10.1016/S0076-6879(01)40444-7

Jasin M (1996) Genetic manipulation of genomes with rare-cutting endonucleases. Trends Genet 12:224–228. doi:10.1016/0168-9525(96)10019-6

Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6:3901–3907

Kempin SA, Liljegren SJ, Block LM, Rounsley SD, Yanofsky MF, Lam E (1997) Targeted disruption in *Arabidopsis*. Nature 389:802–803. doi:10.1038/39770

Kim YG, Cha J, Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci USA 93:1156–1160. doi:10.1073/pnas.93.3.1156

Klug A (2005) The discovery of zinc fingers and their development for practical applications in gene regulation. Proc Jpn Acad 81:87–102. doi:10.2183/pjab.81.87

Kumar S, Fladung M (2001) Controlling transgene integration in plants. Trends Plant Sci 6:155–159. doi:10.1016/S1360-1385(01)01890-8

Lee KY, Lund P, Lowe K, Dunsmuir P (1990) Homologous recombination in plant cells after Agrobacterium-mediated transformation. Plant Cell 2:415–425

Lloyd A, Plaisier CL, Carroll D, Drews GN (2005) Targeted mutagenesis using zinc-finger nucleases in *Arabidopsis*. Proc Natl Acad Sci USA 102:2232–2237. doi:10.1073/pnas.0409339102

Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA, Ando D, Urnov FD, Galli C, Gregory PD, Holmes MC, Naldini L (2007) Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol 25:1298–1306. doi:10.1038/nbt1353

Maeder ML, Thbodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M, Jaing T, Foley JE, Winfrey RJ, Townsend JA, Unger-Wallance E, Sander JD, Muller-Lerch F, Fu F, Pearlberg J, Gobel C, Dassie JP, Pruett-Miller SM, Porteus MH, Sgroi DC, Iafrate AJ, Dobbs D, McCray PB, Cathomen T, Voytas DF, Joung JK (2008) Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell 31:294–301. doi:10.1016/j.molcel.2008.06.016

Miao ZH, Lam E (1995) Targeted disruption of the TGA3 locus in *Arabidopsis thaliana*. Plant J 7:359–365. doi:10.1046/j.1365-313X.1995.7020359.x

Moehle EA, Rock JM, Lee YL, Jouvenot Y, Dekelver RC, Gregory PD, Urnov FD, Holmes MC (2007) Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. Proc Natl Acad Sci USA 104:3055–3060. doi:10.1073/pnas.0611478104

Offringa R, de Groot MJ, Haagsman HJ, Does MP, van den Elzen PJ, Hooykaas PJ (1990) Extrachromosomal homologous recombination and gene targeting in plant cells after Agrobacterium mediated transformation. EMBO J 9:3077–3084

Paszkowski J, Baur M, Bogucki A, Potrykus I (1988) Gene targeting in plants. EMBO J 7:4021–4026

Pavletich NP, Pabo CO (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science 252:809–817. doi:10.1126/science.2028256

Porteus MH, Carroll D (2005) Gene targeting using zinc finger nucleases. Nat Biotechnol 23:967–973

Puchta H (2002) Gene replacement by homologous recombination in plants. Plant Mol Biol 48:173–182

Puchta H (1998) Repair of genomic double-strand breaks in somatic plant cells by one-sided invasion of homologous sequences. Plant J 13:331–339. doi:10.1046/j.1365-313X.1998.00035.x

Puchta H (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. J Exp Bot 56:1–14. doi:10.1093/jxb/eri123

Puchta H, Dujon B, Hohn B (1993) Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. Nucleic Acids Res 21:5034–5040. doi:10.1093/nar/21.22.5034

Reiss B (2003) Homologous recombination and gene targeting in plant cells. Int Rev Cytol 228:85–139. doi:10.1016/S0074-7696(03)28003-7

Reiss B, Schubert I, Kopchen K, Wendeler E, Schell J, Puchta H (2000) RecA stimulates sister chromatid exchange and the fidelity of double-strand break repair, but not gene targeting, in plants transformed by *Agrobacterium*. Proc Natl Acad Sci USA 97:3358–3363. doi:10.1073/pnas.050582797

Shalev G, Sitrit Y, Avivi-Ragolski N, Lichtenstein C, Levy AA (1999) Stimulation of homologous recombination in plants by expression of the bacterial resolvase ruvC. Proc Natl Acad Sci USA 96:7398–7402. doi:10.1073/pnas.96.13.7398

Siebert R, Puchta H (2002) Efficient repair of genomic double-strand breaks by homologous recombination between directly repeated sequences in the plant genome. Plant Cell 14:1121–1131. doi:10.1105/tpc.001727

Terada R, Urawa H, Inagaki Y, Tsugane K, Iida S (2002) Efficient gene targeting by homologous recombination in rice. Nat Biotechnol 20:1030–1034. doi:10.1038/nbt737

Terada R, Johzuka-Hisatomi Y, Saitoh M, Asao H, Iida S (2007) Gene targeting by homologous recombination as a biotechnological tool for rice functional genomics. Plant Physiol 144:846–856. doi:10.1104/pp.107.095992

Thompson W, Hall G, Spiker S, Allen G (1997) A nuclear scaffold attachment region which increases gene expression. WO 97/27207

Trouiller B, Charlot F, Choinard S, Schaefer DG, Nogue F (2007) Comparison of gene targeting efficiencies in two mosses suggests that it is a conserved feature of bryophyte transformation. Biotechnol Lett 29:1591–1598. doi:10.1007/s10529-007-9423-5

Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435:646–651. doi:10.1038/nature03556

Verdaguer B, de Kochko A, Beachy RN, Fauquet C (1996) Isolation and expression in transgenic tobacco and rice plants, of the cassava vein mosaic virus (CVMV) promoter. Plant Mol Biol 31:1129–1139. doi:10.1007/BF00040830

Waldron C, Murphy E, Roberts J, Gustafson G, Armour S, Malcolm S (1985) Resistance to hygromycin B: a new marker for plant transformation studies. Plant Mol Biol 5:103–108. doi:10.1007/BF00020092

Wang HX, Viret JF, Eldridge A, Perera R, Signer ER, Chiurazzi M (2001) Positive-negative selection for homologous recombination in *Arabidopsis*. Gene 272:249–255. doi:10.1016/S0378-1119(01)00532-7

Wohlleben W, Arnold W, Broer I, Hillemann D, Strauch E, Puhler A (1988) Nucleotide sequence of the phosphinothricin *N*-acetyltransferase gene from Streptomyces viridochromogenes Tu494 and its expression in *Nicotiana tabacum*. Gene 70:25–37. doi:10.1016/0378-1119(88)90101-1

Wright DA, Townsend JA, Winfrey RJ Jr, Irwin PA, Rajagopal J, Lonosky PM, Hall BD, Jondle MD, Voytas DF (2005) High-frequency homologous recombination in plants mediated by zinc-finger nucleases. Plant J 44:693–705. doi:10.1111/j.1365-313X.2005.02551.x