High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases

Feng Zhang[a], Morgan L. Maeder[b,c], Erica Unger-Wallace[d], Justin P. Hoshaw[a], Deepak Reyon[d,e], Michelle Christian[a], Xiaohong Li[a], Christopher J. Pierick[a], Drena Dobbs[d], Thomas Peterson[d], J. Keith Joung[b,c,f], and Daniel F. Voytas[a,1]

[a]Department of Genetics, Cell Biology, and Development and Center for Genome Engineering, University of Minnesota, Minneapolis, MN 55455; [b]Molecular Pathology Unit and Center for Cancer Research, Massachusetts General Hospital, Charlestown, MA 02129; [c]Center for Computational and Integrative Biology, Massachusetts General Hospital, Boston, MA 02114; [d]Department of Genetics, Development and Cell Biology, and [e]Interdepartmental Graduate Program in Bioinformatics and Computational Biology, Iowa State University, Ames, IA 50011; and [f]Department of Pathology, Harvard Medical School, Boston, MA 02115

Edited* by Joseph R. Ecker, Salk Intstitute, La Jolla, CA, and approved March 31, 2010 (received for review December 30, 2009)

We report here an efficient method for targeted mutagenesis of *Arabidopsis* genes through regulated expression of zinc finger nucleases (ZFNs)—enzymes engineered to create DNA double-strand breaks at specific target loci. ZFNs recognizing the *Arabidopsis ADH1* and *TT4* genes were made by Oligomerized Pool ENgineering (OPEN)—a publicly available, selection-based platform that yields high quality zinc finger arrays. The *ADH1* and *TT4* ZFNs were placed under control of an estrogen-inducible promoter and introduced into *Arabidopsis* plants by floral-dip transformation. Primary transgenic *Arabidopsis* seedlings induced to express the *ADH1* or *TT4* ZFNs exhibited somatic mutation frequencies of 7% or 16%, respectively. The induced mutations were typically insertions or deletions (1–142 bp) that were localized at the ZFN cleavage site and likely derived from imprecise repair of chromosome breaks by nonhomologous end-joining. Mutations were transmitted to the next generation for 69% of primary transgenics expressing the *ADH1* ZFNs and 33% of transgenics expressing the *TT4* ZFNs. Furthermore, ≈20% of the mutant-producing plants were homozygous for mutations at *ADH1* or *TT4*, indicating that both alleles were disrupted. *ADH1* and *TT4* were chosen as targets for this study because of their selectable or screenable phenotypes (*adh1*, allyl alcohol resistance; *tt4*, lack of anthocyanins in the seed coat). However, the high frequency of observed ZFN-induced mutagenesis suggests that targeted mutations can readily be recovered by simply screening progeny of primary transgenic plants by PCR and DNA sequencing. Taken together, our results suggest that it should now be possible to obtain mutations in any *Arabidopsis* target gene regardless of its mutant phenotype.

nonhomologous end-joining | gene knockout | alcohol dehydrogenase | chalcone synthase

In model organisms such as yeast and mice, targeted mutagenesis is a powerful tool for generating specific DNA sequence alterations that enable a greater understanding of gene function (1–3). Because of its potential for advancing basic plant science, an efficient method for targeted mutagenesis has been a long-sought-after goal of plant biology. Targeted mutagenesis also holds great promise for crop improvement. Precise methods for genome modification will be necessary if crop plants are to be fully harnessed to meet the burgeoning food, fiber, and fuel needs of an expanding world population.

Engineered zinc finger nucleases (ZFNs) are powerful reagents for targeted genome modification (4). ZFNs function as dimers with each monomer composed of an engineered zinc finger array (ZFA), typically consisting of three or four fingers fused to a nonspecific cleavage domain of the FokI endonuclease (Fig. 1*A*). ZFAs can be engineered to bind diverse target DNA sequences within a genome of interest. Binding sites for the two ZFAs (each typically 18 or 24 bp in length) are separated by 5–7 bp, and this spacing allows the FokI monomers to dimerize and create a DNA double-strand break (DSB) in the spacer sequence between the half-sites.

Chromosome breaks created by ZFNs are repaired predominantly by two DNA repair pathways—nonhomologous end-joining (NHEJ) and homologous recombination (HR) (5). Each pathway provides different opportunities for targeted mutagenesis. HR repairs chromosome breaks by copying information from homologous DNA templates; if a DNA template is delivered to cells along with the ZFN, sequence modifications in the template are introduced into the repaired chromosome to create the desired genetic variants. NHEJ also creates genetically useful sequence variation. Faithful repair by NHEJ restores DNA integrity by rejoining the broken chromosomes precisely; however, imprecise repair introduces small insertions/deletions at ZFN cleavage sites that frequently knock out gene function.

Arabidopsis thaliana is one of the premier models for plant biology, and an efficient targeted mutagenesis method would further enhance this plant's utility for experimental biology. Previous work has shown that ZFNs can generate mutations in *Arabidopsis* by NHEJ; for example, a well-characterized ZFN (QQR) was shown to create heritable mutations at an integrated, chromosomal reporter gene with a QQR recognition site (6). HR at a ZFN cleavage site in *Arabidopsis* was also reported (7). The HR experiments did not demonstrate alteration of an endogenous *Arabidopsis* gene, but also used a previously characterized ZFN and an integrated reporter gene with appropriate ZFN recognition sites. Further, the ZFNs used in this study (7) did not stimulate HR much above levels observed in the absence of a targeted chromosome break (i.e., 1 in 3,000 transformed plants) (8). Both studies are important because they demonstrate that ZFNs can be used to modify sequences in *Arabidopsis* by NHEJ and HR. They also identify the challenges in implementing efficient ZFN-induced mutagenesis in this species: (*i*) robust ZFN engineering methods are required that enable endogenous *Arabidopsis* sequences to be targeted, and (*ii*) efficient protocols are needed that allow recovery of germinal NHEJ- and HR-induced mutations at frequencies exceeding 10%, rates previously achieved with ZFNs in *Drosophila* (9), human cells (10), rat (11), zebrafish (12, 13), maize (14), and tobacco (15).

In this study, we report a highly efficient method for targeted mutagenesis of *Arabidopsis* genes by imprecise repair of ZFN-induced chromosome breaks by NHEJ. ZFNs that recognize the *Arabidopsis ADH1* and *TT4* genes were generated by using

Author contributions: F.Z., E.U.-W., D.D., T.P., J.K.J., and D.F.V. designed research; F.Z., M.L.M., J.P.H., D.R., M.C., X.L., and J.K.J. performed research; F.Z., M.L.M., J.P.H., D.R., M.C., C.J.P., D.D., and J.K.J. contributed new reagents/analytic tools; F.Z., M.L.M., J.P.H., J.K.J., and D.F.V. analyzed data; and F.Z. and D.F.V. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

See commentary on page 11657.

[1]To whom correspondence should be addressed. E-mail: voytas@umn.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.0914991107/-/DCSupplemental.

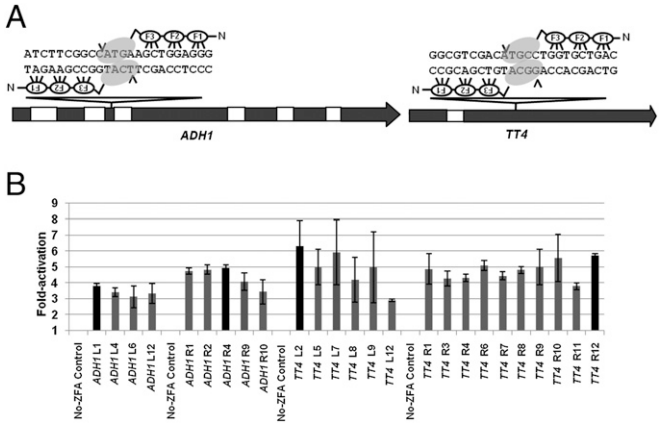

Fig. 1. OPEN ZFN target sites in the *ADH1* and *TT4* coding sequences. (*A*) The gene models of *Arabidopsis ADH1* and *TT4* loci are shown as block arrows. Black rectangles represent exons; white rectangles represent introns. The positions of OPEN ZFN target sites are indicated by triangles. Each engineered ZFA consists of three zinc fingers (F1, F2, F3), which together recognize a 9-bp target site. The two target sites are separated by 6-bp spacer sequences. Binding of the ZFAs to the target sequences enables the FokI nuclease monomers (indicated as gray ovals) to dimerize and cleave within the spacer. The sites of cleavage in the target sequences are indicated by black arrowheads. (*B*) ZFAs generated by OPEN were tested for their activity in bacterial two-hybrid (B2H) assays. The ZFAs were tested for their ability to bind target sequences in bacteria and activate transcription of a downstream *lacZ* reporter gene. Fold activation of each ZFA relative to its negative control, which does not express the ZFA, is plotted on the *y* axis. Multiple ZFAs generated by OPEN were tested for each of the *ADH1* and *TT4* left and right target sites. ZFAs with activities indicated by black bars were used in subsequent tests. Error bars denote SD; $n = 3$.

a publicly available platform (Oligomerized Pool ENgineering, OPEN) that is sufficiently robust to target most *Arabidopsis* genes (16, 17). To make our approach more accessible, we developed a web-based genome browser that identifies potential OPEN ZFN target sites in the *Arabidopsis* genome. Primary transgenic *Arabidopsis* plants expressing OPEN-derived ZFNs early in development generated loss-of-function mutations in the next generation at frequencies ranging from 33 to 69% for *TT4* and *ADH1*, respectively. The efficiency of the mutagenesis indicates that OPEN ZFNs can be routinely and reliably used for targeted genetic modification of *Arabidopsis*.

Results

Engineering ZFAs That Recognize Endogenous *Arabidopsis* Genes. We used OPEN (17) to generate ZFAs for potential target sites in two *Arabidopsis* genes: *ALCOHOL DEHYDROGENASE1* (*ADH1*; AT1G77120) and *TRANSPARENT TESTA4* (*TT4*; AT5G13930) (Fig. 1*A*). The web-based program, ZiFiT (18) was used to identify OPEN target sites within approximately the first one-third of the ORF to increase the likelihood that ZFN-induced mutations would result in loss of gene function. Four OPEN selections (one for each half-site of the two full ZFN target sites) were performed, and 12 resulting ZFAs for each half-site were tested for activity in a bacterial two-hybrid assay (17). Between four and nine ZFAs per half-site activated the expression of a *lacZ* reporter gene ≈3-fold or more over negative controls (Fig. 1*B* and Table S1). Previous studies have shown that activities >3-fold reliably identify ZFAs with a high probability of functioning as ZFNs in eukaryotic cells (16). A single ZFA with the highest activity for each half-site was chosen for subsequent tests.

Testing OPEN ZFNs in Yeast and *Arabidopsis* Protoplasts. The *ADH1* and *TT4* ZFAs were fused to FokI and tested in a yeast-based assay for their ability to function as ZFNs. This assay uses a target plasmid and two ZFN expression plasmids that are brought together in the same cell by mating (Fig. 2*A*) (15). The target plasmid has a *lacZ* reporter gene with a 125-bp duplication of coding sequence. The duplication flanks a *URA3* gene and a target site recognized by a given ZFN. A ZFN-induced DSB at the target site is repaired through single-strand annealing between the duplicated sequences, creating a functional *lacZ* gene and resulting in loss of *URA3*. Using this assay, the ZFAs targeting the *ADH1* and *TT4* genes were found to function effectively as ZFNs (Fig. 2*B*). LacZ activity resulting from cleavage by the *ADH1* ZFN pair and subsequent recombination-based repair was comparable with activity observed for a control ZFN engineered from the well-characterized ZFA, Zif268 (19). Activity of the *TT4* ZFN pair was ≈2-fold less than that of the Zif268 ZFN.

ZFN pairs were next tested for their ability to recognize and cleave native chromosomal sites in *Arabidopsis* protoplasts. ZFNs were cloned into expression vectors containing the 35S promoter, and the respective pairs of *ADH1* and *TT4* ZFNs were introduced into *Arabidopsis* protoplasts by PEG-mediated transformation (20). Cleavage activity was assessed at 24 or 48 h after transformation by the presence of target site mutations introduced through imprecise repair by NHEJ. To detect such mutations, we adopted a previously described PCR assay that preferentially amplifies mutated DNA sequences (6). In this assay, genomic DNAs extracted from transformed protoplasts are first digested with restriction enzymes (NlaIII for *ADH1* and NspI for *TT4*), which occur within the ZFN cleavage site (Fig. 3*A*). Because the restriction sites are destroyed by most NHEJ-induced mutations,

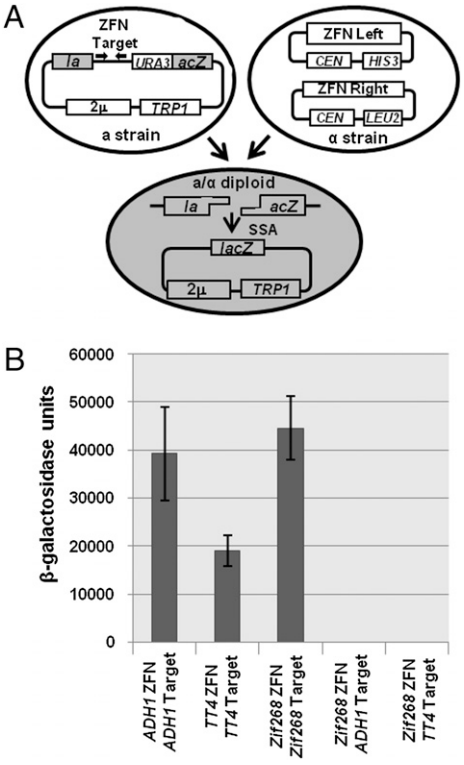

Fig. 2. ZFN activity in yeast. (*A*) Schematic of the yeast-based ZFN activity assay. ZFN expression plasmids are introduced into a yeast strain of α mating-type; the reporter plasmid is transformed into a yeast strain of **a** mating type. Mating of the two strains brings together the ZFNs and the reporter plasmid into a diploid cell. This results in cleavage of the ZFN target sequence and restoration of the *lacZ* gene through single strand annealing (SSA). (*B*) Results of the yeast assay performed with the *ADH1* and *TT4* ZFNs. A ZFN derived from the well-characterized ZFA, Zif268, is included as the positive control. Negative controls are the *ADH1* and *TT4* target sites combined with the Zif268 ZFN. Error bars denote SD; $n = 3$.

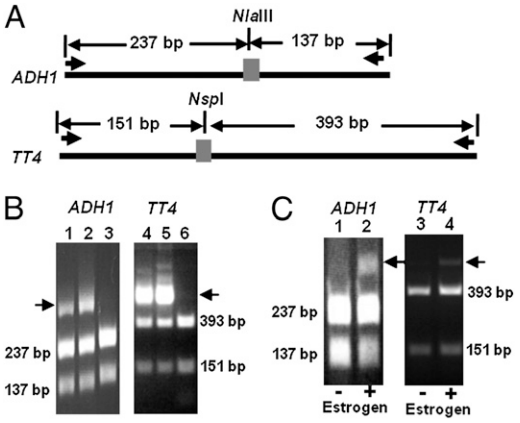

Fig. 3. Detection of ZFN-induced mutations in *Arabidopsis* protoplasts and somatic cells. (*A*) Schematic of the strategy used to amplify and identify mutations at the ZFN target sites (gray rectangle). Restriction enzyme sites coincident with the ZFN cleavage sites are NlaIII for *ADH1* and NspI for *TT4*. Short black arrows indicate the positions of PCR primers relative to the target sites. (*B*) Restriction endonuclease assay to detect ZFN-induced mutations in *Arabidopsis* protoplasts. Mutations introduced by NHEJ are resistant to restriction enzyme digestion due to loss of restriction sites and result in uncleaved PCR products (indicated by arrows). Genomic DNA was cleaved with restriction enzymes before PCR amplification to enrich for ZFN-induced mutations. Lanes: 1, protoplast sample transformed with *ADH1* ZFN after 24 h incubation; 2, protoplast sample transformed with *ADH1* ZFN after 48 h of incubation; 3, control protoplast sample without transformation; 4, protoplast sample transformed with *TT4* ZFN after 24 h of incubation; 5, protoplast sample transformed with *TT4* ZFN after 48 h of incubation; 6, control protoplast sample without transformation; (*C*) Restriction endonuclease assay to detect ZFN-induced mutations in *Arabidopsis* seedlings after estrogen induction. DNA fragments lacking restriction sites (arrows) are detected in experimental samples with estrogen induction but not in control samples without estrogen. In this case, enrichment by prior restriction digestion of genomic DNA was not performed. Lanes: 1, *ADH1* ZFN transgenic T_1 plants without estrogen induction; 2, *ADH1* ZFN transgenic T_1 plants with estrogen induction; 3, *TT4* ZFN transgenic T_1 plants without estrogen induction; 4, *TT4* ZFN transgenic T_1 plants with estrogen induction.

the mutated sequences are resistant to restriction digestion and are amplified preferentially in a subsequent round of PCR with primers flanking the target sites. When PCR products are further digested by the same restriction enzymes and subjected to gel electrophoresis, uncleaved sequences with mutations can be distinguished from cleaved wild-type sequences. Using this assay, mutated alleles were observed in DNA samples derived from cells transformed with either *ADH1* or *TT4* ZFN pairs (Fig. 3*B*).

ZFN-induced mutations in the uncleaved PCR products were verified and characterized by cloning and DNA sequencing. Forty-eight clones lacking NlaIII or NspI sites (24 clones each) were sequenced from samples transformed with either the *ADH1* or *TT4* ZFNs. DNA sequences of all clones contained mutations at cognate target sites. Eighteen of 24 clones from the *ADH1* ZFN treatment and 16 of 24 clones from the *TT4* treatment showed distinct mutations, indicating that they were independent events. Mutations included both insertions (from 1 to 58 bp) and deletions (from 2 to 53 bp). All insertions >1 bp consisted of DNA sequences that are also present in the ZFN expression vectors (Fig. S1); extrachromosomal plasmid DNA was observed to serve as filler DNA in the repair of DNA DSBs in tobacco cells (21). In conclusion, ZFN pairs generated by OPEN for *Arabidopsis ADH1* and *TT4* loci induce mutations at target sites in their native chromosomal context.

High Frequency ZFN-Induced Mutagenesis in *Arabidopsis* Somatic Tissues. The ability of OPEN ZFNs to induce mutations was next tested in planta. *ADH1* and *TT4* ZFN pairs were introduced into wild-type *Arabidopsis* plants by the *Agrobacterium* floral dip method (22). An estrogen-inducible expression system was used to regulate production of the ZFNs (23, 24). After floral dip transformation, seeds (T_0) were collected and germinated on Murashige–Skoog (MS) plates that contained hygromycin and 17β-estradiol. Ten days after germination, the primary transgenic plants (T_1) were resistant to hygromycin, and the expression of ZFNs in these T_1 plants should have been induced by the 17β-estradiol. Mutations at *ADH1* and *TT4* target sites were assessed by limited-cycle PCR using DNA prepared from pooled seedlings (in this case with no prior restriction digestion). The PCRs were digested with NlaIII or NspI, and the presence of undigested PCR products suggested that both the *ADH1* and *TT4* ZFNs induced mutations in the seedlings (Fig. 3*C*). No undigested PCR products were observed in uninduced plants. To confirm the mutagenesis and to estimate mutagenesis frequencies, the PCR products from induced transgenic plants were cloned and sequenced. In the T_1 seedlings with the *ADH1* ZFN, 16% (14/90) of the clones contained distinct mutations at the target site, whereas in the induced *TT4* ZFN T_1 seedlings, 7% (7/95) of the clones showed distinct target site mutations (Table 1). The identified mutations included insertions (1–2 bp), deletions (3–142 bp), and one case of a single nucleotide substitution derived from the *TT4* ZFN (Fig. S2). Taken together, these results indicate that expression of OPEN ZFNs for both *ADH1* and *TT4* genes efficiently creates mutations at their target sites in *Arabidopsis* somatic tissues.

Cleavage of ZFNs at noncognate sites can generate unwanted background mutations (15, 25). To look for evidence of off-target cleavage, we examined DNA sequences in the *Arabidopsis* genome that were most similar to the target sites of the *ADH1* and *TT4* ZFNs. The sequence most similar to the *ADH1* target site has one mismatch in the left half-site and two mismatches in the right half-site, whereas the sequence most similar to the *TT4* target site has one mismatch in each half-site (Table S2). PCR amplification was carried out on the DNA samples used above to detect somatic mutations. No evidence of ZFN-induced mutations was detected by sequencing 96 clones derived from PCR products that span the potential off-target sequences. The observed preference of OPEN ZFNs for their intended targets indicates that these reagents have a high degree of specificity in vivo.

Efficient Germinal Transmission of ZFN-Induced Mutations. Mutations occurring early in development in the L2 cells of the shoot apical meristem are likely to be transmitted to the next generation (26). To determine whether ZFN-induced mutations at *ADH1* could be transmitted germinally, 16 T_1 plants exposed to estrogen during germination were grown to maturity in soil, and seeds were collected from individual lines. Fifty seeds from each line were then subjected to allyl alcohol treatment. A functional ADH1 gene product converts allyl alcohol to highly toxic acrylaldehyde, so wild-type *ADH1* seeds will die after treatment, whereas mutant *adh1* seeds will survive (Fig. S3*A*) (27). As shown in Table 1, 69% (11 of 16) of the induced T_1 plants give rise to *adh1* mutants. Nine of the 11 plants segregated mutant and wild-type phenotypes in their progeny (Table 1). Two plants (18%) produced exclusively *adh1* mutants, indicating that both *ADH1* alleles were disrupted (i.e., biallelic modification). The

Table 1. Frequencies of ZFN-induced mutations at *ADH1* and *TT4*

ZFN	Somatic mutation frequency, %	Frequency of mutant-producing T1 plants	Percent frame shift/in-frame mutant alleles, %
ADH1 ZFN	16	69% (11/16)	73/27
TT4 ZFN	7	33% (10/30)	88/12

mutations likely occurred early in development and gave rise to homozygous mutant floral tissue. Sequence of the *ADH1* target site in all mutant plants uncovered 11 distinct mutations, all of which were insertions/deletions at the ZFN recognition site (Fig. 4*A*). Furthermore, three mutant plants, including the plant with exclusively *adh1* progeny, gave rise to distinct mutation types (Fig. 4*A*), providing molecular evidence that independent mutations can occur in both alleles of the *ADH1* gene.

The germinal transmission frequency of ZFN-induced mutations at *TT4* was determined by scoring progeny derived from 30 estrogen-exposed T_1 transgenic plants. Disruption of *TT4* results in colorless seed coats (28), but because the seed coat is derived from maternal tissues, only homozygous mutant flowers, in which both alleles sustained an independent mutation, give rise to colorless seeds (Fig. S3*B*). We observed that seven of 30 T_1 plants produced colorless seeds (≈20–30 mutant seeds of 1,000), indicating that those plants contained clones of cells with both *TT4* alleles disrupted. Biallelic disruption was corroborated by PCR amplification and DNA sequencing of the ZFN cleavage site in the T_2 plants derived from the homozygous mutant seeds (Fig. 4*B*). In addition, as shown in Fig. 4*B*, three of seven mutant-producing T_1 plants contained independent mutations in both *TT4* alleles. Because monoallelic gene disruption cannot be scored by the seed coat color phenotype, T_2 progeny derived from the remaining T_1 plants were subjected to a PCR screen. PCR amplification was carried out on genomic DNA made from 20 pooled seedlings derived from each line. Digestion with NspI identified an additional three T_1 plants that produced heritable *tt4* mutant DNA sequences. In agreement with the molecular analyses, we also observed that the T_2 progeny of these three T_1 plants segregated *tt4* homozygous mutants, which lack purple anthocyanin pigmentation in the cotyledon and hypocotyl (29). The mutations in these T_2 plants were verified by DNA sequencing (Fig. 4*B*). Taken together, the germinal mutagenesis frequency for *TT4* was estimated to be 33% (10/30; Table 1).

A *ADH1*

```
CACGTATCTTCGGCCATGAAGCTGGAGGGTAATA
---------------ATGAAGCTGGAGGGTAATA  -51
CACGTATCTTCGGCCA--AAGCTGGAGGGTAATA  -2[1]
------------------------GAGGGTAATA  -28
CACGTATCTTCGGCCAT--AGCTGGAGGGTAATA  -2[2]
CACGTATCTTCGGCCA---AGCTGGAGGGTAATA  -3[3]
CACGTATCTTCGGCC---------GAGGGTAATA  -9[1]
CACGTATCTTCGGCCA----GCTGGAGGGTAATA  -4
CACGTATCTTCGGCCAaTGAAGCTGGAGGGTAAT  +1[3]
CACGTATCTTCGGCCAaTGAgGCTGGAGGGTAAT  +1
CACGTATCTTCGGCCAtaTGAAGCTGGAGGGTAA  +2[2]
CACGTATCTTCGGCCAtgaTGAAGCTGGAGGGTA  +3
```

B *TT4*

```
CCTCCGGCGTCGACATGCCTGGTGCTGACTACCA
-----------------CCTGGTGCTGACTACCA  -61[1]
---------------------GTGCTGACTACCA  -65[1]
CCTCCGGCGTCGACAT---TGGTGCTGACTACCA  -3[3]
-------------------TGGTGCTGACTACCA  -65
CCTCCGGCGTCGACA-----GGTGCTGACTACCA  -5
CCTCCGGCGTCGACATtGCCTGGTGCTGACTACC  +1[2]
CCTCCGGCGTCGACATGcCCTGGTGCTGACTACC  +1
CCTCCGGCGTCGACATgtGCCTGGTGCTGACTAC  +2[2,3]
```

Fig. 4. Sequences of germinally transmitted mutations induced by ZFNs. For each target gene, the wild-type sequence is shown at the top with the ZFN recognition sites underlined. The recovered mutant alleles are shown below the wild-type sequence. Deletions are indicated by dashed lines, and insertions are indicated by lowercase letters. (*A*) Germinal mutations at *ADH1*. The numbers in superscript (i.e., 1, 2 and 3) identify mutations derived from individual T_1 plants. In many cases, a given plant produced multiple independent mutations. For example, plant 1 produced exclusively *adh1* double knock-out (i.e., biallelic) mutant progeny, and the two distinct mutant alleles are indicated by the superscript 1. (*B*) Germinal mutations at *TT4*. The numbers in superscript identify T1 plants that contain multiple independent mutations at both alleles of *TT4*.

Web-Based Tool To Identify ZFN Target Sites in the *Arabidopsis* Genome. To aid plant biologists in performing ZFN-mediated mutagenesis of *Arabidopsis*, ZiFiT was used to identify all potential OPEN ZFN target sites in the *Arabidopsis* genome (18). In total, 381,497 ZFN target sites were identified in 76% (30,193 of 39,640) of *Arabidopsis* gene models (TAIR9 release). Of these sites, 171,409 sites are located in protein coding regions (mean of ≈4.3 sites per gene model). A web-based genome browser, ZFNGenome, was implemented to display OPEN ZFN sites along with gene annotations (Fig. S4). ZFNGenome is freely available and can be accessed at http://bindr.gdcb.iastate.edu:8888/ZFNGenome/Arabidopsis/.

Discussion

We report here a highly efficient method for targeted mutagenesis of *Arabidopsis* genes. ZFNs were used to create chromosome breaks at specific, endogenous loci, and the cleaved chromosomes were repaired imprecisely by nonhomologous end-joining to yield small DNA insertions or deletions that disrupted gene function. Previous reports on the use of ZFNs for targeted mutagenesis in *Arabidopsis* used ZFNs derived from ZFAs that do not recognize endogenous *Arabidopsis* sequences (6, 7). Rather, reporter constructs with the appropriate ZFN recognition sites were first introduced into the *Arabidopsis* genome, and these integrated reporters were then used to measure targeted mutagenesis with the corresponding ZFN. We believe the key to making endogenous sequences readily amenable to ZFN-induced modification was the development of the OPEN platform for ZFN engineering (16). To date, OPEN has been used to engineer ZFNs that recognize 16 different target sequences in the human (16, 30), zebrafish (31), tobacco (15), and now *Arabidopsis* genomes. Further, as its name implies, OPEN is an open-source platform, and we anticipate that its continued practice will lead to improvements that streamline this approach to ZFN engineering, thereby making it even more accessible to plant biologists.

To aid *Arabidopsis* researchers in using OPEN, we developed a genome browser (ZFNGenome) that displays all possible OPEN ZFN target sites along with gene model annotations. The density of ZFN sites in a target sequence is determined by the reagents available for OPEN selections, which enable engineering of ZFAs composed of all GNN and some TNN subsites (16). Even with available reagents, OPEN ZFN sites are present in more than three-quarters of the *Arabidopsis* coding sequences, and the average gene has more than four different OPEN ZFN sites. As additional reagents for OPEN selections are developed, the number of sites that can be targeted will grow considerably. For example, availability of reagents for ANN subsites would enable 130 OPEN ZFNs to be generated per kb of DNA (formula for calculation in ref. 16).

In addition to the use of OPEN, a second key to the success of our targeted mutagenesis strategy was our approach to limit ZFN toxicity. Toxicity has been reported in human cells, zebrafish, and *Drosophila* and is likely due to activity of ZFNs at off-target sites (12, 13, 31–33). Reducing toxicity in these systems was achieved by increasing ZFN specificity (e.g., by using OPEN) and/or by modifying the FokI nuclease domain to make it function as an obligate heterodimer (16, 34, 35). This latter approach ensures that DNA is cleaved only when ZFNs for the left and right half-sites are positioned at the appropriate target sequence on the chromosome. Limited expression of ZFNs either by transient DNA delivery, injected RNA, or inducible expression systems has also been suggested as a strategy to reduce undesired toxic effects (32, 33, 36).

Toxicity was of particular concern in our targeted mutagenesis strategy, because in contrast to most mutagenesis approaches with ZFNs, we created transgenic *Arabidopsis* plants that have the ZFN expression construct stably integrated. Transgenic

plants were also created in the first report of the use of a ZFN in *Arabidopsis* (6), and transgenic *Drosophila* were generated in initial reports of targeted mutagenesis in this species (9). In both the initial *Drosophila* and *Arabidopsis* studies, heat shock-inducible expression systems were used to control ZFN expression (6, 33). A second, recent report of a ZFN stably integrated into the *Arabidopsis* genome did not describe toxicity when the ZFN was expressed constitutively (7); however, in our experiments, we did not recover *Arabidopsis* transformants with certain ZFNs, suggesting that their expression was deleterious. To minimize toxicity, we used the obligate heterodimeric form of the FokI nuclease in all in planta experiments (34). Further, we chose an estrogen-inducible system (23) to regulate expression because (*i*) relative to heat shock, estrogen-induced expression is more tightly controlled with no background expression observed (37); (*ii*) induction with 17β-estradiol is rapid (i.e., within 30 min) and expression reaches levels 8-fold higher than those attained with the constitutive 35S promoter (23); and (*iii*) 17β-estradiol does not affect *Arabidopsis*, whereas heat pulses can be stressful to young seedlings and flowers (38). These features allowed us to induce ZFN expression at early developmental stages (i.e., during seed germination) and then turn off expression to reduce potential toxicity associated with prolonged exposure of the *Arabidopsis* genome to ZFNs.

The estrogen-inducible system proved highly effective, allowing us to recover mutations in seeds generated from 33 to 69% of the primary transgenic plants. Using a heat shock promoter to drive expression of the QQR ZFN, Lloyd et al. (6) reported mutations in progeny of 10% of the heat-induced T_1 plants. We initially tested a similar protocol of heat-induction by using the *ADH1* and *TT4* ZFNs under control of an *Arabidopsis* heat shock promoter; however, somatic ZFN-induced mutations could be detected only if they were enriched before PCR by first cleaving the genomic DNA with the restriction enzyme that recognizes the spacer sequence between the two ZFN half-sites. Furthermore, no germinal events were recovered in the progeny of >200 T_1 plants that were exposed to the heat shock regime. Although we cannot explain why our attempt with heat induction failed, we believe the very high burst of ZFN expression afforded by estrogen induction was critical to success. The resulting efficiency of mutagenesis made it possible to look at small populations of T_1 plants (<30) and, therefore, it should be possible to recover mutations in genes without observable phenotypes by simply screening progeny of primary transgenic plants by PCR and DNA sequencing.

Although all T_1 plants generated in this study by using OPEN ZFN constructs looked normal and healthy, there is still a possibility of off-target cleavage, which has been reported and could create unwanted secondary mutations (15, 25). Off-target cleavage is consistent with our observation that stable transgenics could not be recovered with some constitutively expressed ZFNs. In this study, examination of DNA sequences in the *Arabidopsis* genome most similar to the target sites of *ADH1* and *TT4* ZFNs provided no evidence of off-target cleavage. Although we cannot rule out off-target activity in such a small survey, any unlinked off-target mutations that do arise with the use of ZFNs can be removed by a simple backcross, which is standard practice in other mutagenesis protocols with *Arabidopsis*.

The spectrum of mutations generated with ZFNs through imprecise nonhomologous end-joining offers much opportunity for genetic analysis. The independent alleles generated by ZFNs are not limited to null mutants that cause frame shifts in the coding sequences, but some are alleles with in-frame insertions/deletions (i.e., the size of the insertion/deletion is a multiple of 3). For instance, three *adh1* alleles and one *tt4* allele identified in this study showed in-frame deletions or insertions (Fig. 4). Such a wide spectrum of allelic variants will provide valuable opportunities to better understand gene function.

The *Arabidopsis* community is fortunate to have large collections of mutants generated by T-DNA. However, in one of the largest T-DNA mutant collections generated at the Salk Institute (39), ≈20% of genes do not have T-DNA insertions and many T-DNA insertions do not provide null phenotypes. Furthermore, 17% of the genes in *Arabidopsis* exist as tandem arrays, so even when two T-DNA lines exist with insertions in each member of a tandem array, it is very difficult to recover double mutants by genetic recombination, because the two genes are so tightly linked (40). ZFNs could make this fraction of the genome amenable to genetic analysis. For example, a recent report (41) suggests it is possible to use ZFNs to create large deletions of genomic DNA at high efficiencies. In summary, we believe targeted mutagenesis with ZFNs is now practical in *Arabidopsis* and is a valuable addition to the suite of genetic tools available for this important plant model.

Materials and Methods

OPEN Selections and ZFN Expression Vectors. OPEN selections and the bacterial two-hybrid assay were conducted as described (Fig. S5) (16, 17). The resulting ZFAs were PCR amplified from phagemids (primers OK.1677 and OK.1678 (31); Table S3), digested with XbaI and BamHI, and cloned into yeast expression vectors pCP3 and pCP4, which have *HIS3* and *LEU2* marker genes, respectively. Each vector also encodes a homodimeric FokI nuclease domain with the SV40 nuclear localization signal (NLS); expression is driven by the yeast *TEF1* promoter. The *ADH1* and *TT4* ZFAs for the left half-sites were cloned into pCP3 (generating pFZ9 and pFZ11, respectively); ZFAs for the right half-sites were cloned into pCP4 (generating pFZ10 and pFZ12, respectively). The plant expression vectors, pFZ14 and pFZ15, were used for the protoplast assays and encode left and right obligate heterodimeric FokI domains (34) with the SV40 NLS; expression is driven by the 35S promoter. ZFAs for the left half-sites were cloned into pFZ14 (generating pJH1, *ADH1*; pJH9, *TT4*); ZFAs for right half-sites were cloned into pFZ15 (generating pJH2, *ADH1*; pJH10, *TT4*).

An estrogen-inducible T-DNA expression vector, pMDC7 (23, 24), was used to generate transgenic *Arabidopsis* plants. As a Gateway-compatible destination vector, pMDC7 can recombine with an entry vector that contains ZFAs fused with FokI nuclease. The *ADH1* and *TT4* ZFN entry vectors encode the left and right ZFNs. Between the ZFN coding sequences is an in-frame 18-amino acid ribosome skipping signal (T2A), which allows translation of multiple ORFs from a single transcript (42). The final *ADH1* and *TT4* ZFN expression vectors (pFZ62 and pFZ63, respectively) were made by recombining the entry vectors with pMDC7 (Fig. S6).

Testing ZFNs in Yeast and Plants. The yeast assay and the assay for testing ZFN activity in protoplasts were carried out as described (15, 20). To generate stably transformed *Arabidopsis* lines, the T-DNA expression vectors pFZ62 and pFZ63 were transformed into *Agrobacterium* strain GV3101 (43). Floral-dip transformation was conducted by using ecotype Columbia (22). Transformed seeds were plated onto solid Murashige and Skoog (MS) medium containing 25 mg/L hygromycin to select for transformants and 10 μM 17β-estadiol to induce expression of ZFNs. The seedlings were grown under a regime of 16 h light/8 h dark at 21 °C in a growth chamber. To evaluate somatic mutations, genomic DNA was prepared from 10 7-day-old T_1 seedlings. High fidelity PCR was then performed to amplify target sites (primers oFZ1–oFZ4; Table S3). PCR products were TOPO cloned (Invitrogen), and 96 clones were randomly selected for DNA sequencing.

To evaluate germinal mutations, hygromycin resistant seedlings (T_1) were transferred to soil at 10 days and grown to maturity. To screen for *adh1* mutants, seeds were collected from 16 transgenic plants carrying the *ADH1* ZFN constructs. Approximately 50 seeds from each T_1 plant were treated with allyl alcohol (27) and grown on MS medium plates. Homozygous *adh1* mutants typically survive allyl alcohol treatment, and genomic DNA from eight surviving sibling plants were pooled and subjected to high fidelity PCR (primers, oFZ1 and oFZ2; Table S3). PCR products were cloned and sequenced as described above.

The *tt4* mutants were identified by screening 30 T_1 transgenic plants carrying *TT4* ZFN constructs to identify those that produced seeds with colorless seed coats. For those plants producing putative *tt4* mutants, DNA was prepared from eight mutant seedlings and genotyped by using high fidelity PCR (primers, oFZ3 and oFZ4; Table S3). PCR products were cloned and sequenced as described above. For plants that did not show mutant phenotypes, DNA was prepared from ≈20 seedlings, and the ZFN cleav-

age site was amplified by PCR. PCR products were subjected to NspI digestion; mutant sequences can be readily identified as being resistant to restriction enzyme digestion. The undigested DNA molecules were TOPO cloned and sequenced.

Sequences similar to the target sites of *ADH1* and *TT4* ZFNs were identifed by using a web-based pattern search tool (44). Two sequences (one for each target site) most similar to the target sites were chosen for subsequent genotyping. The region encompassing the candidate cleavage site was PCR amplified (primers oFZ5–oFZ8; Table S3), and the products were TOPO cloned. DNA sequences were obtained from 96 independent clones from each site.

ZFNGenome. *Arabidopsis* chromosomal contigs were scanned for potential OPEN target sites by using a variation of the ZiFiT algorithm (18). Potential target sites were limited to those for which ZFNs can be engineered by using currently available OPEN reagents (17). Only protein coding regions were searched by using gene models from TAIR9 release (45). Spacers between the two half-sites of 5–7 nts were considered. Sites that contain potential *dam* or *dcm* methylation sites were excluded, as were sites without a GNN subsite, because previous studies have shown that successful OPEN sites have at least one GNN (17). A web browser (ZFNGenome) was implemented to display all potential OPEN sites along with gene annotations. The browser used a Berkeley DB backend (46) linked to a Generic Genome Browser (GBrowse) frontend via open source adaptors from BioPerl (47). ZFN target sites were included along with default features, such as "loci" and "gene models." Each ZFN target site is hyperlinked to ZiFDB and ZiFiT (18, 48).

ACKNOWLEDGMENTS. We thank David Wright for assistance in ZFN design, Yi Hou for help with the Arabidopsis protoplast assay, and Kenichi Tsuda for help with plasmid construction. The National Science Foundation supported work carried out by F.Z., J.P.H., M.C., and D.F.V. (MCB 0209818) and D.R. and D.D. (DBI 0923827). M.L.M. and J.K.J. are supported by National Institutes of Health Grants R01 GM069906, R24 GM078369, and R21 RR024189 and the Massachusetts General Hospital Pathology Service.

1. Doetschman T, et al. (1987) Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature* 330:576–578.
2. Rothstein RJ (1983) One-step gene disruption in yeast. *Methods Enzymol* 101: 202–211.
3. Thomas KR, Capecchi MR (1987) Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* 51:503–512.
4. Cathomen T, Joung JK (2008) Zinc-finger nucleases: The next generation emerges. *Mol Ther* 16:1200–1207.
5. Wyman C, Kanaar R (2006) DNA double-strand break repair: All's well that ends well. *Annu Rev Genet* 40:363–383.
6. 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.
7. de Pater S, Neuteboom LW, Pinas JE, Hooykaas PJ, van der Zaal BJ (2009) ZFN-induced mutagenesis and gene-targeting in Arabidopsis through Agrobacterium-mediated floral dip transformation. *Plant Biotechnol J* 7:821–835.
8. Britt AB, May GD (2003) Re-engineering plant gene targeting. *Trends Plant Sci* 8: 90–95.
9. Bibikova M, Beumer K, Trautman JK, Carroll D (2003) Enhancing gene targeting with designed zinc finger nucleases. *Science* 300:764.
10. Urnov FD, et al. (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* 435:646–651.
11. Geurts AM, et al. (2009) Knockout rats via embryo microinjection of zinc-finger nucleases. *Science* 325:433.
12. Doyon Y, et al. (2008) Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat Biotechnol* 26:702–708.
13. Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA (2008) Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat Biotechnol* 26:695–701.
14. Shukla VK, et al. (2009) Precise genome modification in the crop species Zea mays using zinc-finger nucleases. *Nature* 459:437–441.
15. Townsend JA, et al. (2009) High-frequency modification of plant genes using engineered zinc-finger nucleases. *Nature* 459:442–445.
16. Maeder ML, et al. (2008) Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol Cell* 31:294–301.
17. Maeder ML, Thibodeau-Beganny S, Sander JD, Voytas DF, Joung JK (2009) Oligomerized pool engineering (OPEN): an 'open-source' protocol for making customized zinc-finger arrays. *Nat Protoc* 4:1471–1501.
18. Sander JD, Zaback P, Joung JK, Voytas DF, Dobbs D (2007) Zinc Finger Targeter (ZiFiT): An engineered zinc finger/target site design tool. *Nucleic Acids Res* 35:W599–W605.
19. Pavletich NP, Pabo CO (1991) Zinc finger-DNA recognition: Crystal structure of a Zif268-DNA complex at 2.1 A. *Science* 252:809–817.
20. Hoshaw JP, Unger-Wallace E, Zhang F, Voytas DF (2010) A transient assay for monitoring zinc-finger nuclease activity at endogenous plant gene targets. *Methods Mol Biol*, in press.
21. Gorbunova V, Levy AA (1997) Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. *Nucleic Acids Res* 25:4650–4657.
22. Clough SJ, Bent AF (1998) Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* 16:735–743.
23. Zuo J, Niu QW, Chua NH (2000) Technical advance: An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *Plant J* 24:265–273.
24. Curtis MD, Grossniklaus U (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol* 133:462–469.
25. Perez EE, et al. (2008) Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol* 26:808–816.
26. Irish VF, Jenik PD (2001) Cell lineage, cell signaling and the control of plant morphogenesis. *Curr Opin Genet Dev* 11:424–430.
27. Jacobs M, Dolferus R, Van den Bossche D (1988) Isolation and biochemical analysis of ethyl methanesulfonate-induced alcohol dehydrogenase null mutants of arabidopsis thaliana (L.) Heynh. *Biochem Genet* 26:105–122.
28. Li J, Ou-Lee TM, Raba R, Amundson RG, Last RL (1993) Arabidopsis flavonoid mutants are hypersensitive to UV-B irradiation. *Plant Cell* 5:171–179.
29. Shirley BW, et al. (1995) Analysis of Arabidopsis mutants deficient in flavonoid biosynthesis. *Plant J* 8:659–671.
30. Zou J, et al. (2009) Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell Stem Cell* 5:97–110.
31. Foley JE, et al. (2009) Rapid mutation of endogenous zebrafish genes using zinc finger nucleases made by Oligomerized Pool ENgineering (OPEN). *PLoS One* 4:e4348.
32. Pruett-Miller SM, Reading DW, Porter SN, Porteus MH (2009) Attenuation of zinc finger nuclease toxicity by small-molecule regulation of protein levels. *PLoS Genet* 5: e1000376.
33. Bibikova M, Golic M, Golic KG, Carroll D (2002) Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics* 161:1169–1175.
34. Miller JC, et al. (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol* 25:778–785.
35. Szczepek M, et al. (2007) Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol* 25:786–793.
36. Beumer KJ, et al. (2008) Efficient gene targeting in Drosophila by direct embryo injection with zinc-finger nucleases. *Proc Natl Acad Sci USA* 105:19821–19826.
37. Masclaux F, Charpenteau M, Takahashi T, Pont-Lezica R, Galaud JP (2004) Gene silencing using a heat-inducible RNAi system in Arabidopsis. *Biochem Biophys Res Commun* 321:364–369.
38. Kim SY, Hong CB, Lee I (2001) Heat shock stress causes stage-specific male sterility in Arabidopsis thaliana. *J Plant Res* 114:301–307.
39. Alonso JM, et al. (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science* 301:653–657.
40. Jander G, Barth C (2007) Tandem gene arrays: A challenge for functional genomics. *Trends Plant Sci* 12:203–210.
41. Lee HJ, Kim E, Kim JS (2010) Targeted chromosomal deletions in human cells using zinc finger nucleases. *Genome Res* 20:81–89.
42. Halpin C, Cooke SE, Barakate A, El Amrani A, Ryan MD (1999) Self-processing 2A-polyproteins–a system for co-ordinate expression of multiple proteins in transgenic plants. *Plant J* 17:453–459.
43. Koncz C, Schell J (1986) The promoter of the TL-DNA gene 5 controls the tissue-specific expression of chimeric genes carried by a novel type of Agrobacterium binary vector. *Mol Gen Genet* 204:383–396.
44. Dong Q, et al. (2005) Comparative plant genomics resources at PlantGDB. *Plant Physiol* 139:610–618.
45. Swarbreck D, et al. (2008) The Arabidopsis Information Resource (TAIR): Gene structure and function annotation. *Nucleic Acids Res* 36:D1009–D1014.
46. Olson M, Bostic K, Seltzer M, Berkeley D (1999) Proceedings of the annual conference on USENIX Annual Technical Conference.
47. Stajich JE, et al. (2002) The Bioperl toolkit: Perl modules for the life sciences. *Genome Res* 12:1611–1618.
48. Fu F, et al. (2009) Zinc Finger Database (ZiFDB): A repository for information on C2H2 zinc fingers and engineered zinc-finger arrays. *Nucleic Acids Res* 37:D279–D283.