Review TRENDS in Plant Science Vol.12 No.5



# Tandem gene arrays: a challenge for functional genomics

Georg Jander[1] and Carina Barth[2]

[1] Boyce Thompson Institute for Plant Research, Tower Road, Cornell University, Ithaca, NY 14853, USA
[2] Department of Biology, 53 Campus Drive, West Virginia University, Morgantown, WV 26506, USA

**In sequenced plant genomes, 15% or more of the identified genes are members of tandem-arrayed gene families. Because mutating only one gene in a duplicated pair often produces no measurable phenotype, this poses a particular challenge for functional analysis. To generate phenotypic knockouts, it is necessary to create deletions that affect multiple genes, select for rare meiotic recombination between tightly linked loci, or perform sequential mutant screens in the same plant line. Successfully implemented strategies include PCR-based screening for fast neutron-induced deletions, selection for recombination between herbicide resistance markers, and localized transposon mutagenesis. Here, we review the relative merits of current genetic approaches and discuss the prospect of site-directed mutagenesis for generating elusive knockouts of tandem-arrayed gene families.**

## Tandem gene multiplications pose a challenge for functional analysis

Completed (*Arabidopsis thaliana*, rice and poplar), in-progress (tomato, *Medicago truncatula* and maize), and pending plant genome projects will provide an unprecedented amount of information about the protein repertoire of higher plants. However, with estimates of the number of protein-coding regions ranging from 27 000 in *Arabidopsis* to 59 000 in maize [1–3], functional genetic analysis will be one of the greatest challenges facing plant biologists in coming years. Even in *Arabidopsis*, currently the best-studied model plant, there is direct experimental evidence for the function of only about a third of the predicted genes [2].

One of the more conclusive approaches for assigning function to unknown plant genes is to study phenotypic changes associated with knockout mutations. In several model plant species, efficient reverse genetics approaches, including TILLING [4], transposon mutagenesis [5–14], and large collections of sequenced T-DNA insertions [15–19], enable the identification of knockout mutations in almost any gene of interest. However, functional analysis is complicated by the lack of measurable phenotypes associated with most knockout mutations [20]. A major contributing factor to this observation is the highly redundant nature of plant genomes. It is likely that all known plant species have polyploid origins [21,22]. In *Arabidopsis*, the most recent genome duplication event occurred ~40 million years ago (mya) whereas the diploid maize probably evolved from an allotetraploid origin within the past 5 my [23]. By contrast, the hexaploid wheat genome is the result of polyploidization events that occurred during the domestication process [24]. Another significant contribution to genome redundancy comes from tandem gene multiplications, where two or more homologous genes are directly adjacent to one another in the genome. Although enumeration of tandem arrays is influenced by the particular bioinformatic analyses that are employed, it is clear that they represent a significant proportion of all plant genes: 17% in *Arabidopsis* [1], 14% in rice [25], 16% in poplar [26], and 35% in maize [3]. Most tandem multiplications in *Arabidopsis* (75%) and rice (79%) contain only two genes, and gene arrays with more than three members are uncommon [27].

Because of genetic redundancy, it is often necessary to knock out more than one member of a plant gene family to obtain a mutant phenotype [28–31]. Creating such double mutants can be relatively simple if the genes of interest are genetically unlinked and non-essential: mutations are identified in two separate plant lines, the mutant lines are crossed (e.g. a/a B/B × A/A b/b), and 1/16 of the resulting $F_2$ progeny are homozygous for both mutations. However, in the absence of reliable site-directed mutagenesis protocols for higher plants, finding double mutants of tightly linked genes is difficult. For instance, if two mutations are 5000 bp apart in the *Arabidopsis* genome, then genetic recombination would bring both mutations together in one out of every 8000 gametes (assuming 200 kbp/cM in *Arabidopsis*). Homozygous double mutants would be found as roughly one out of every 64 million $F_2$ progeny or one out of every 32 000 $F_3$ progeny from such a cross. Thus, except in rare cases where there is a simple screen or selection for the desired phenotype, finding such double mutants can be challenging using standard genetic methods. Here we describe alternate strategies (Table 1) that can be used to identify knockouts of two or more tandem-arrayed plant genes for functional analysis.

## Selection for deletions spanning multiple genes

Large chromosomal deletions, unlike point mutations and DNA insertions, readily knock out the function of multiple adjacent genes. Therefore, strategies for the efficient generation and detection of chromosomal deletions will facilitate the functional analysis of tandem-arrayed genes. Treatment with a 60 Gy fast neutron dose creates about ten deletions per genome, knocking out individual *Arabidopsis*

*Corresponding author:* Jander, G. (gj32@cornell.edu).
Available online 9 April 2007.

 1360-1385/$ – see front matter  doi:10.1016/j.tplants.2007.03.008

**Table 1. Advantages and disadvantages of approaches for generating mutations in tandem-arrayed genes**

| Approach | Advantages | Disadvantages | Refs |
|---|---|---|---|
| Selection for genomic deletions | • Can delete multiple adjacent genes<br>• Mutagenesis is performed on dry seeds<br>• No plant transformation is required<br>• Effective, particularly for targeting small genes<br>• Fairly fast owing to multiplexed PCR screens | • Requires DNA sequence data for region adjacent to the genes of interest<br>• Can delete nearby genes, resulting in phenotypes not related to gene of interest<br>• Requires generation of a large mutant population | [33,34] |
| Localized *Ds* transposition and excision by Cre-*loxP* | • *Arabidopsis Ds-loxP* T-DNA insertion resource is being developed<br>• Can delete multiple adjacent genes<br>• Does not require full genome sequence<br>• Effective PCR strategies allow screening for gene excision | • Requires two crosses to generate a deletion<br>• Resource is not available for most plant species<br>• Requires plant transformation to generate *Ds-loxP* T-DNA insertion population | [5,35] |
| RNA interference | • Allows downregulation of multiple genes using one construct<br>• Variable silencing efficiency is useful if null mutations are lethal<br>• Does not require a full genome sequence | • Possible off-target effects<br>• Requires transformation and screening in subsequent generations<br>• Effectiveness varies, depending on DNA constructs, integration site and target genes | [38–44] |
| miRNA-based gene silencing | • Allows silencing of multiple genes<br>• miRNA design software is available<br>• Relatively fast and easy to implement<br>• Does not require full genome sequence | • Degree of silencing can vary when targeting multiple genes<br>• Possible off-target effects of miRNAs<br>• Requires plant transformation<br>• Degree of silencing can vary when targeting multiple genes | [45,46] |
| Selection for meiotic recombination | • Selection is based on herbicide resistance<br>• Homozygous lethal double mutants can be identified in the heterozygous state<br>• Insertion mutant populations are available for *Arabidopsis* and rice<br>• Can select complete gene knockouts | • Depends on rare meiotic recombination<br>• Requires many crosses and PCR screens<br>• Tedious in the absence of insertion mutants<br>• Gene knockout populations are not available for most plant species<br>• Not good for multi-gene families | [51,53] |
| Sequential mutagenesis | • Heterozygous transposon insertions can be detected by PCR<br>• Does not require a full genome sequence<br>• Transposon launch sites are available for several model plant species<br>• Can select complete gene knockouts | • Can require a large mutant population for reverse genetics screening (e.g. by TILLING)<br>• Can be a time- and labor-intensive approach<br>• Transposon mutagenesis requires crosses<br>• Not good for multi-gene families | [61] |
| Site-directed mutagenesis by zinc finger nucleases (ZFNs) | • Highly specific when using appropriately designed zinc finger domains<br>• Might allow deletion of multiple genes<br>• Software for identification of potential ZFN target sites is available<br>• Automated zinc finger design methods are being developed | • Not yet a proven technology<br>• Limited by the ability to engineer zinc finger domains that recognize specific targets<br>• Possible off-target effects<br>• Potential cytotoxicity of ZFNs | [70–74] |

genes with a frequency of one in 2500 [32]. For detecting such fast neutron-generated deletions, PCR primer extension times can be adjusted such that the deletion allele of a given locus is selectively amplified (Figure 1a). Xin Li *et al.* [33] pioneered this amplification strategy by showing that it is sensitive enough to detect chromosomal deletions in a background of 2500-fold excess non-mutant DNA. By screening successively smaller pools from a population of >50 000 fast neutron-mutagenized *Arabidopsis* lines, the same group identified deletions affecting 84% of the targeted loci (21 of 25 tested). These included the tandem-duplicated *TGA2* and *TGA6* transcription factors [34], as well as a three-gene array [33]. This deletion-detection approach has also been successfully employed with rice [33] and, in theory, could be applied to any plant species. However, it depends on having several thousand base pairs of DNA sequence information flanking the locus of interest.

An alternate approach for generating chromosomal deletions makes use of localized transposition by the maize *Dissociation* (*Ds*) element and site-specific DNA excision by the bacteriophage Cre-*loxP* recombination system [35]. First, a *Ds-loxP* T-DNA construct, containing *loxP* sites in both the *Ds* element and the T-DNA backbone is transformed into the genome (Figure 1b). Next, transposition of *Ds-loxP* is induced by crossing in *Activator* (*Ac*). Typically, 10% of all reinsertions are within 10 kbp of the *Ds-loxP* T-DNA insertion site [35]. Finally, after a second cross to bring in the *Cre* recombinase gene, Cre-mediated site-specific recombination between the two *loxP* sites causes the excision of the intervening DNA. Excision of a gene of interest can be detected by PCR strategies such as those described above for fast neutron-induced deletions. A community resource of >10 000 sequenced *Ds-loxP* T-DNA insertion sites has been generated for *Arabidopsis* researchers [5]. This collection, which will eventually be expanded to 50 000 lines (P. Krysan, personal communication), will contain *Ds-loxP* T-DNA insertions throughout the genome and will allow the targeted deletion of many, perhaps most *Arabidopsis* tandem-arrayed gene families. One disadvantage to the current strategy is the necessity to perform two crosses to generate a deletion, first to bring in *Ac* and next to bring in *Cre*. However, use of a heat shock-inducible *Ac* transposase [7] in future *Ds-loxP* constructs will eliminate at

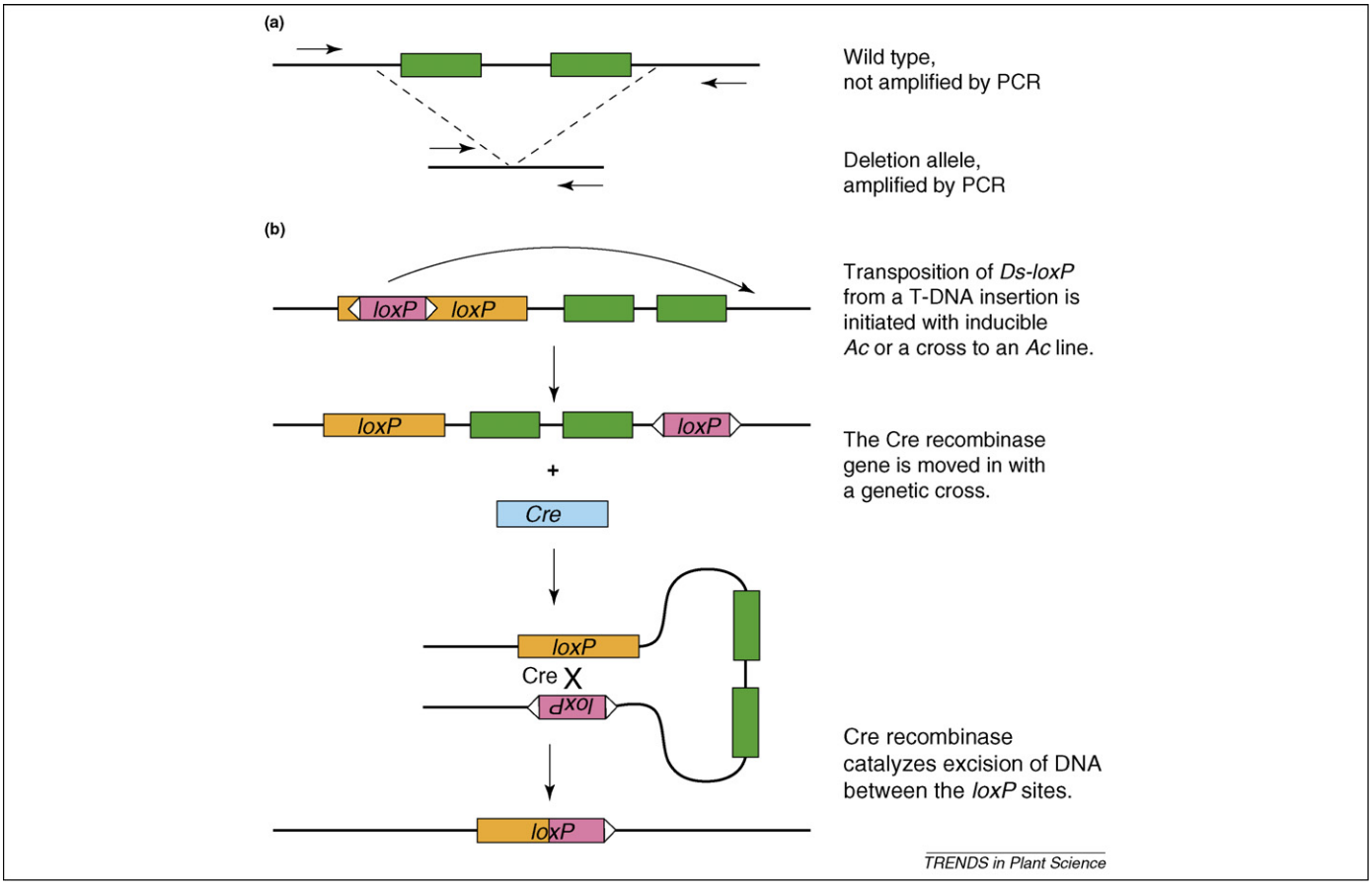


**Figure 1**. Deletion of tandem-arrayed genes. **(a)** Fast neutron-induced mutations can be detected in pooled DNA samples by adjusting PCR conditions so that only the shorter deletion allele can be amplified [33]. PCR primers are indicated as arrows. **(b)** Deletion of DNA flanking *Ds-loxP* T-DNA insertions is accomplished by inducing transposition of the *Ds-loxP*, bringing in Cre recombinase with a genetic cross, and screening for Cre-mediated deletion of the intervening DNA between the *loxP* sites in the *Ds* element and the T-DNA [35].

least the first cross. *Ac-Ds* transposons are active in many plant species and, in theory, could be broadly applicable for generating deletions. A similar recombinase system, employing R/*RS* from the yeast *Zygosaccharomyces rouxii* in combination with *Ac*, has been shown to induce deletion formation in rice [36] and petunia [37].

## Silencing gene families by RNA interference

RNA interference, which allows down-regulation of gene expression in transgenic plants, is an attractive method for functional analysis of gene families, including those where the genes of interest are located in tandem arrays. Gene silencing by RNA interference results from sequence-specific RNA degradation that follows the formation of double-stranded RNA that is homologous in sequence to the targeted gene [38–40]. By targeting shared sequences, it is possible to reduce expression of multiple members of a gene family simultaneously. For instance, seven members of the *OsRac* gene family in rice were silenced with varying efficiency by a single RNA interference construct [41]. In hexaploid wheat (*Triticum aestivum*), silencing of homeologous genes with double-stranded RNA constructs induced phenotypes that would have been difficult or impossible to attain by standard genetic approaches [42].

Rather than targeting multigene families with a single shared RNA segment, it is possible to achieve effective suppression using chimeric silencing constructs made with pieces of several genes. For instance, a chimeric hairpin RNA construct was used to silence all members of the opium poppy (*Papaver somniferum*) codeinone reductase gene family [43]. Similarly, a chimeric construct employing partial sequences of several genes was used to suppress lignin biosynthetic enzymes in tobacco (*Nicotiana tabacum*) [44].

The use of synthetic microRNA constructs represents an alternate approach to using relatively large RNA fragments for gene silencing. Two recent studies demonstrated that down-regulation of gene families by endogenous and synthetic microRNAs mimics the phenotypes produced by previously described multiple mutations in the targeted genes [45,46]. Software for automated design of artificial microRNAs, based on the backbone of a natural microRNA and the input of individual or several related target sequences, is available (http://wmd.weigelworld.org).

The RNA-based gene silencing approaches described above can easily be applied to reduce the expression of duplicated genes. However, interpretation of the results can be complicated by the variable effectiveness of gene silencing by RNA interference. Silencing efficiency depends on the type of construct that is used, the site of integration, and the particular target genes [47–50]. Furthermore, because gene silencing is generally not 100% effective, mutant phenotypes that require the complete knockout of a gene family might not be detected

using the described silencing approaches. For instance, insertional knockout of a single defense-related *Arabidopsis* myrosinase reduced total enzymatic activity by more than 95% (similar to what one might achieve by RNA interference), but had no measurable effect on insect resistance. By contrast, a *tgg1 tgg2* double myrosinase knockout reduced enzymatic activity to undetectable levels and made the plants much more sensitive to lepidopteran herbivory [51].

### Selection for meiotic recombination between adjacent genes

Most tandem arrays in plants contain only two genes [27]. Therefore, if viable mutants are available for both loci, an obvious approach is to cross them and identify homozygous recombinants in subsequent generations. However, as discussed above, in the absence of a strong screen or selection for the expected phenotype, finding meiotic recombination between tightly linked genes is difficult. Screening for recombinants among thousands of progeny from a cross is feasible only for mutations that cause easily scored visible phenotypes, for instance kernel color changes induced by *bronze* mutations [48] and intragenic recombinants at the *r* locus [52] in maize. Although it is theoretically possible to screen several thousand plants by PCR-based methods to identify double mutants, we are not aware of any published examples of this approach.

An alternate approach for identifying rare meiotic recombination between closely linked plant genes relies on mutant alleles that confer herbicide resistance. The first implementation of this method involved selection for intragenic recombination between the *Arabidopsis csr1-1* and *csr1-2* mutations, which confer resistance to chlorsulfuron and imidazolinone herbicides, respectively [53]. First, the two mutants were crossed to create a double-resistant heterozygous line with the mutations in *trans*. Then, the heterozygote was used as the male parent in a second cross, and meiotic recombination bringing the mutations in *cis* was identified by transfer of both herbicide resistance genes in one pollen grain. This genetic strategy can be generalized to select for meiotic recombination between any two *Arabidopsis* insertional mutations that confer resistance to different herbicides (Figure 2). Meiotic recombination between T-DNA insertions conferring Basta and kanamycin resistance, respectively, in the tandem-duplicated *TGG1* and *TGG2 Arabidopsis* myrosinase genes was found with a frequency of one in 3000 [51]. Use of a male sterile mutant as the pollen recipient facilitated the otherwise laborious number of crosses that were needed to find the desired recombinants.

An added advantage of the approach outlined in Figure 2 is that double mutants are selected in the heterozygous state. This reduces the number of plants that need to be screened to identify recombinants and also allows

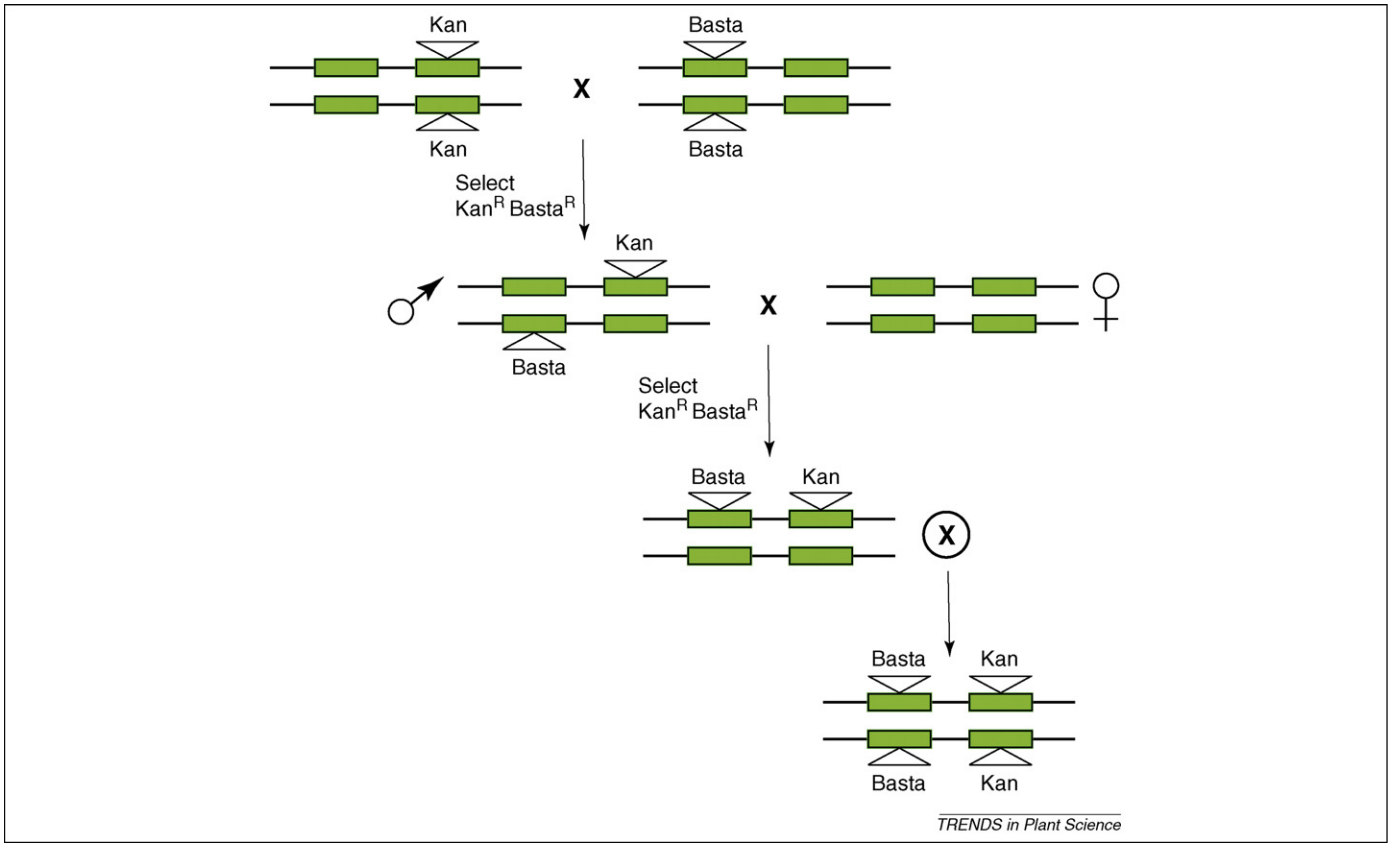


**Figure 2**. Selection for recombination between T-DNA or transposon insertions [51]. Insertions conferring herbicide resistance (in this example kanamycin$^R$ and Basta$^R$) are identified in two genes of interest. The first cross brings insertions in tandem-duplicated genes into a *trans* configuration. Progeny of the first cross are used as the pollen donors for a second cross, preferably to a male sterile recipient to reduce the amount of effort involved. Transfer of both herbicide resistance markers in the same pollen grain selects for meiotic recombination that brings the insertions into the *cis* configuration. Self-pollination of these plants allows screening for homozygous double mutants.

identification of homozygous lethal double mutations. In this case, self-pollination of the doubly heterozygous individuals would lead to seed or seedling lethality in ~25% of the progeny. Given the large collections of T-DNA and transposon insertion mutants that are available to *Arabidopsis* researchers [15–17,54,55], it is likely that the approach outlined in Figure 2 can be applied to many other tandem-duplicated gene pairs. Large collections of insertional mutants are also being developed for rice [18,19,56–60], making a similar approach feasible for this important crop plant.

## Sequential mutagenesis to create a double mutant

Rather than identifying two mutations independently and bringing them together by genetic recombination, it is also possible to select for a second mutation in the genetic background of the first one. If the second mutation causes a known, easily scored phenotype, this could be fairly straightforward and would essentially represent a genetic enhancer screen. Otherwise, reverse genetics approaches such as TILLING [4] could be used to identify a second mutation without regard to the possible phenotype. However, the hurdles involved in setting up a large mutant population and using reverse genetics to find one particular mutated gene can be prohibitive. We are aware of no published reports where such approaches have been applied to generate double knockouts of tandem-duplicated plant genes.

Available collections of sequenced transposon launch sites in *Arabidopsis*, maize, rice and tomato [5–14] allow localized saturation mutagenesis of plant genomes. As was demonstrated recently with the *Arabidopsis CYP79F1* and *CYP79F2* genes [61], the propensity of *Ac–Ds* to both reinsert nearby in the genome and leave a 'footprint' in the DNA sequence at the original insertion site can be used to create double mutants of adjacent genes (Figure 3). First, a *Ds* transposon insertion in *CYP79F2* was reactivated by induction of transposase expressed from a heat shock promoter. Next transposition into the adjacent *CYP79F1* gene was identified by visually screening for a known *cyp79F1* rosette phenotype in the second generation after transposase induction. Finally the transposon footprint in *cyp79F2* and the reinsertion in *cyp79F1* were verified by PCR amplification and DNA sequencing. In theory, this localized transposon mutagenesis could be used to identify double mutants of many other genes. However, not all such selections will benefit from easily identified phenotypes such as those associated with the *cyp79F1*. In most cases, it will be necessary to screen for the desired second insertion using PCR-based approaches. However, PCR-based screening has the advantage that heterozygous insertional mutations can be detected in the first generation after transposon mutagenesis. By contrast, visually identifying recessive *cyp79F1* homozygotes required screening hundreds of mutant families in the second generation after mutagenesis [61]. Pooling strategies, where multiple DNA samples are combined for PCR and individual lines are analyzed only if a pool containing those lines produces a positive result [6], can greatly reduce the number of PCR reactions required to find transposon insertions in a particular gene.

## Prospects for site-directed mutagenesis in higher plants

All the methods described above would be trivialized by the development of an effective site-directed mutagenesis protocol for plants. Although targeted mutagenesis is commonly used to study model organisms such as yeast and mice, there is as yet no working system for creating such mutations in higher plants. To date, high frequency homologous recombination has been achieved only in the moss *Physcomitrella patens* [62,63]. In flowering plants, oligonucleotide-directed gene targeting resulting in selectable phenotypes has been demonstrated in tobacco [64] and maize [65,66]. However, gene conversion occurred with a frequency of only $10^{-4}$ in these experiments [64,66]. Gene targeting by homologous recombination in *Arabidopsis* was enhanced more than tenfold by expression of the yeast (*Saccharomyces cerevisiae*) *RAD54* gene [67], a member of the *SWI2/SNF2* superfamily of chromatin remodeling genes [68].

Zinc finger nucleases (ZFNs) represent an exciting new approach for site-directed mutagenesis in plants. These chimeric protein constructs, consisting of sequence-specific

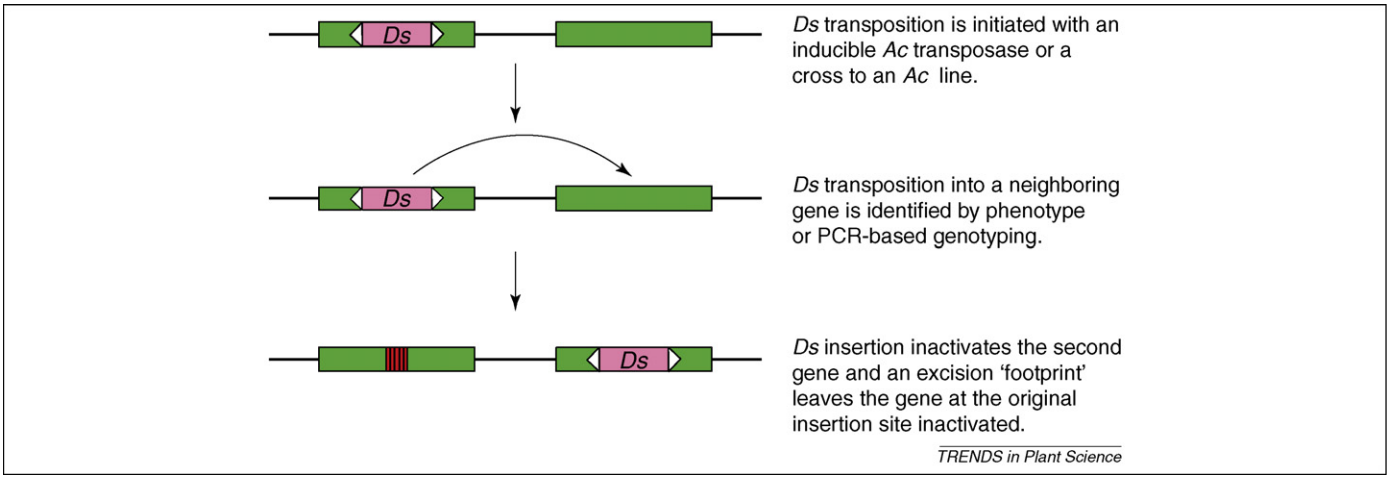


**Figure 3.** Localized transposon mutagenesis. *Ds* elements tend to transpose to nearby genetic locations, allowing screens for insertions in adjacent genes. Imperfect DNA repair can leave the gene at the original insertion site inactivated [61].

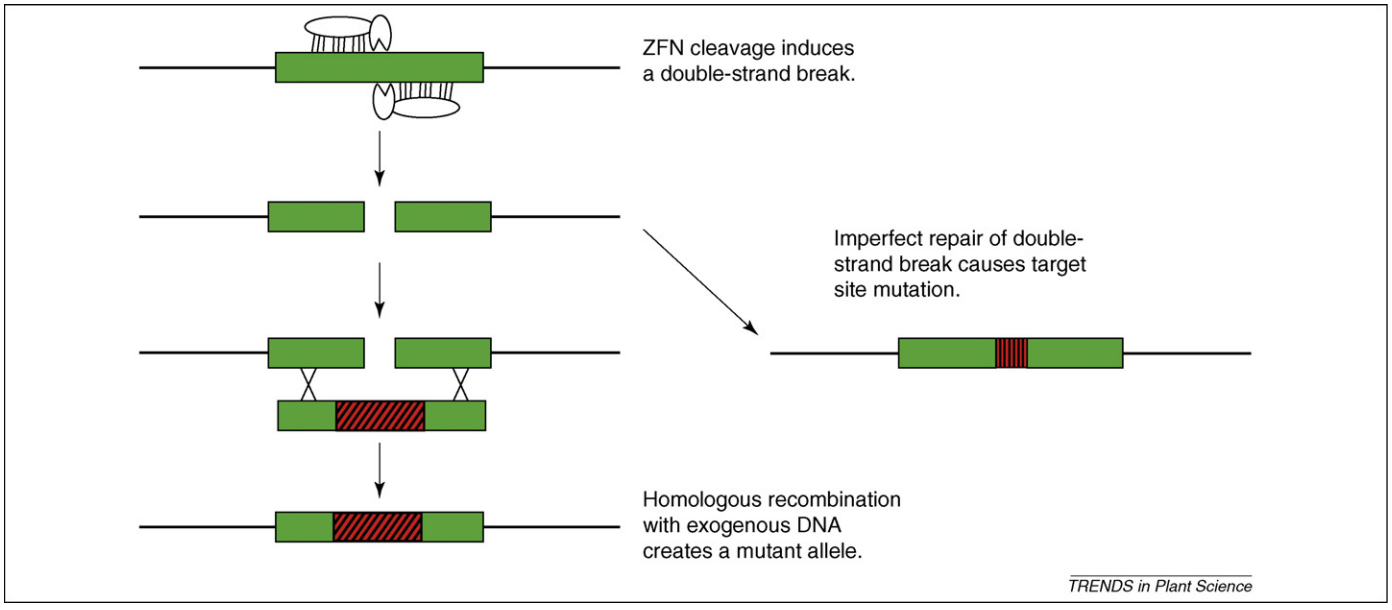


**Figure 4**. Site-directed mutagenesis with a zinc finger nuclease (ZFN). Double-strand cleavage by a ZFN leads to target gene knockout by imperfect repair [70] or homologous recombination [71].

zinc finger DNA binding domains and a *Fok*I DNA nuclease, stimulate site-specific recombination and repair processes by generating double-strand breaks at specific target sites in genomic DNA [69]. Because double-strand breaks stimulate DNA repair and recombination in essentially all organisms, targeted cleavage with ZFNs should be broadly applicable for site-directed mutagenesis. Two proof-of-concept publications [70,71] show that ZFNs can induce chromosomal mutations in *Arabidopsis* (Figure 4). Alan Lloyd *et al.* [70] combined a ZFN gene driven by a heat-shock promoter with a transgenic target in the *Arabidopsis* genome. ZFN cleavage at the specific genomic site stimulated imperfect repair mutations with a frequency of ~10%. In an alternate approach, David Wright *et al.* [71] demonstrated that chromosome breaks created by ZFNs greatly enhance the frequency of homologous recombination. An artificial target gene containing a translational fusion between β-glucuronidase (GUS) and neomycin phosphotransferase (NPTII) was inactivated by a deletion in the GUS active site and the insertion of a ZFN recognition site. ZFN cleavage and restoration of a functional chromosomal *GUS:NPTII* fusion by homologous recombination with transfected DNA occurred with a frequency of ~20%.

Although targeted mutagenesis using ZFNs has great potential, this approach is currently limited by the difficulty of engineering new zinc finger domains, each of which recognizes a specific nucleotide triplet. Modular assembly of individual zinc finger domains has allowed the design of proteins that recognize 18 bp DNA sequences with high specificity and affinity [72]. Given an 18-bp target sequence, it should be possible to identify unique zinc finger DNA binding sites in most plant genes. Not surprisingly, several groups are developing semi-automated methods for designing zinc finger DNA binding proteins that target specific DNA sequences [72–74]. In one such approach [74], the procedure involves the identification of potential ZFN target sites using ZiFiT software (http://bindr.gdcb.iastate.edu/ZiFiT/), assembly of multi-finger arrays, testing the DNA-binding activity of a three-finger domain using a bacterial two-hybrid reporter assay, and construction of ZFN expression vectors. All the required steps take about four weeks. However, the use of ZFNs for homologous recombination *in planta* still requires optimization to avoid off-target DNA breaks and potential cytotoxicity.

With further improvements in the design of ZFNs that target specific genomic sequences, generating knockouts of tandem-arrayed plant genes might become routine. For instance cleavage with a ZFN could be used to induce a second mutation in a tandem-duplicated gene pair in a line where one of the genes already has a T-DNA insertion. Alternatively, if a single ZFN creates double-strand breaks in two members of a gene array, it might be possible to delete the intervening sequence and thereby inactivate multiple genes in one step. In either case, one will be able to use knockout mutations of tandem gene arrays as a tool for their functional analysis.

Although site-directed mutagenesis protocols for higher plants are still being optimized, the other genetic strategies described in this review (summarized in Table 1) are available for making tightly linked double mutations. Together, the described methods represent an effective genetic toolkit for plant researchers. The particular species being investigated and the available mutant alleles will dictate which approach is the best choice for the functional analysis of tandem-arrayed plant genes.

## Acknowledgements

We thank Dan Voytas and Patrick Krysan for providing access to their unpublished manuscripts, and Hong Gu Kang, Tom Brutnell and Martin De Vos for helpful discussions. We apologize to the authors of relevant papers that we were not able to cite because of space constraints.

## References

1 The *Arabidopsis* Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815

2 The Multinational *Arabidopsis* Steering Committee (2006) *The Multinational Coordinated* Arabidopsis thaliana *Functional Genomics Project: Annual Report 2006* (http://www.Arabidopsis.org/portals/masc/2006MASC_report.pdf)
3 Messing, J. *et al.* (2004) Sequence composition and genome organization of maize. *Proc. Natl. Acad. Sci. U. S. A.* 101, 14349–14354
4 McCallum, C.M. *et al.* (2000) Targeting Induced Local Lesions IN Genomes (TILLING) for plant functional genomics. *Plant Physiol.* 123, 439–442
5 Woody, S.T. *et al.* (2007) The *WiscDsLox* T-DNA collection: an *Arabidopsis* community resource generated using an improved high-throughput T-DNA sequencing pipeline. *J. Plant Res.* 120, 157–165
6 Kolkman, J.M. *et al.* (2005) Distribution of *Activator* (*Ac*) throughout the maize genome for use in regional mutagenesis. *Genetics* 169, 981–995
7 Nishal, B. *et al.* (2005) An inducible targeted tagging system for localized saturation mutagenesis in *Arabidopsis*. *Plant Physiol.* 137, 3–12
8 Muskett, P.R. *et al.* (2003) A resource of mapped *Dissociation* launch pads for targeted insertional mutagenesis in the *Arabidopsis* genome. *Plant Physiol.* 132, 506–516
9 Greco, R. *et al.* (2003) Transpositional behaviour of an *Ac*/*Ds* system for reverse genetics in rice. *Theor. Appl. Genet.* 108, 10–24
10 Gidoni, D. *et al.* (2003) Multi-functional T-DNA/*Ds* tomato lines designed for gene cloning and molecular and physical dissection of the tomato genome. *Plant Mol. Biol.* 51, 83–98
11 Raina, S. *et al.* (2002) A collection of sequenced and mapped *Ds* transposon insertion sites in *Arabidopsis thaliana*. *Plant Mol. Biol.* 50, 93–110
12 Upadhyaya, N.M. *et al.* (2006) *Dissociation* (*Ds*) constructs, mapped *Ds* launch pads and a transiently-expressed transposase system suitable for localized insertional mutagenesis in rice. *Theor. Appl. Genet.* 112, 1236–1241
13 Kumar, C.S. *et al.* (2005) Efficient insertional mutagenesis in rice using the maize *En*/*Spm* elements. *Plant J.* 44, 879–892
14 An, G. *et al.* (2005) Reverse genetic approaches for functional genomics of rice. *Plant Mol. Biol.* 59, 111–123
15 Alonso, J.M. *et al.* (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301, 653–657
16 Rosso, M.G. *et al.* (2003) An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Mol. Biol.* 53, 247–259
17 Sessions, A. *et al.* (2002) A high-throughput *Arabidopsis* reverse genetics system. *Plant Cell* 14, 2985–2994
18 Zhang, J. *et al.* (2006) RMD: a rice mutant database for functional analysis of the rice genome. *Nucleic Acids Res.* 34, D745–D748
19 Jeong, D.H. *et al.* (2006) Generation of a flanking sequence-tag database for activation-tagging lines in japonica rice. *Plant J.* 45, 123–132
20 Bouche, N. and Bouchez, D. (2001) *Arabidopsis* gene knockout: phenotypes wanted. *Curr. Opin. Plant Biol.* 4, 111–117
21 Blanc, G. and Wolfe, K.H. (2004) Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell* 16, 1667–1678
22 Lockton, S. and Gaut, B.S. (2005) Plant conserved non-coding sequences and paralogue evolution. *Trends Genet.* 21, 60–65
23 Swigonova, Z. *et al.* (2004) Close split of sorghum and maize genome progenitors. *Genome Res.* 14, 1916–1923
24 Buckler, E.S. *et al.* (2001) Molecular diversity, structure and domestication of grasses. *Genet. Res.* 77, 213–218
25 International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. *Nature* 436, 793–800
26 Tuskan, G.A. *et al.* (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604
27 Rizzon, C. *et al.* (2006) Striking similarities in the genomic distribution of tandemly arrayed genes in *Arabidopsis* and rice. *PLoS Comput. Biol.* 2, 989–1000
28 Sibout, R. *et al.* (2005) *CINNAMYL ALCOHOL DEHYDROGENASE-C* and *-D* are the primary genes involved in lignin biosynthesis in the floral stem of *Arabidopsis*. *Plant Cell* 17, 2059–2076
29 Chaiwongsar, S. *et al.* (2006) The protein kinase genes *MAP3Kε1* and *MAP3Kε2* are required for pollen viability in *Arabidopsis thaliana*. *Plant J.* 48, 193–205
30 Tilton, G.B. *et al.* (2006) Plant coenzyme A biosynthesis: characterization of two pantothenate kinases from *Arabidopsis*. *Plant Mol. Biol.* 61, 629–642
31 Enns, L.C. *et al.* (2005) Two callose synthases, GSL1 and GSL5, play an essential and redundant role in plant and pollen development and in fertility. *Plant Mol. Biol.* 58, 333–349
32 Koornneef, M. *et al.* (1982) EMS- and radiation-induced mutation frequencies at individual loci in *Arabidopsis thaliana* (L.) Heynh. *Mutat. Res.* 93, 109–123
33 Li, X. *et al.* (2001) A fast neutron deletion mutagenesis-based reverse genetics system for plants. *Plant J.* 27, 235–242
34 Zhang, Y. *et al.* (2003) Knockout analysis of *Arabidopsis* transcription factors TGA2, TGA5, and TGA6 reveals their redundant and essential roles in systemic acquired resistance. *Plant Cell* 15, 2647–2653
35 Zhang, S. *et al.* (2003) Resources for targeted insertional and deletional mutagenesis in *Arabidopsis*. *Plant Mol. Biol.* 53, 133–150
36 Nakagawa, Y. *et al.* (2001) A system to induce the deletion of genomic sequences using R/*RS* site-specific recombination and the *Ac* transposon in transgenic rice plants. *Theor. Appl. Genet.* 102, 1136–1141
37 Liu, Z. *et al.* (2000) TREGED: a new strategy for inducing deletions in plant genomes. *Plant Mol. Biol. Rep.* 18, 255–263
38 Baulcombe, D. (2004) RNA silencing in plants. *Nature* 431, 356–363
39 Marx, J. (2000) Interfering with gene expression. *Science* 288, 1370–1372
40 Carthew, R.W. (2001) Gene silencing by double-stranded RNA. *Curr. Opin. Cell Biol.* 13, 244–248
41 Miki, D. *et al.* (2005) RNA silencing of single and multiple members in a gene family of rice. *Plant Physiol.* 138, 1903–1913
42 Travella, S. *et al.* (2006) RNA interference-based gene silencing as an efficient tool for functional genomics in hexaploid bread wheat. *Plant Physiol.* 142, 6–20
43 Allen, R.S. *et al.* (2004) RNAi-mediated replacement of morphine with the nonnarcotic alkaloid reticuline in opium poppy. *Nat. Biotechnol.* 22, 1559–1566
44 Abbott, J.C. *et al.* (2002) Simultaneous suppression of multiple genes by single transgenes. Down-regulation of three unrelated lignin biosynthetic genes in tobacco. *Plant Physiol.* 128, 844–853
45 Alvarez, J.P. *et al.* (2006) Endogenous and synthetic microRNAs stimulate simultaneous, efficient, and localized regulation of multiple targets in diverse species. *Plant Cell* 18, 1134–1151
46 Schwab, R. *et al.* (2006) Highly specific gene silencing by artificial microRNAs in *Arabidopsis*. *Plant Cell* 18, 1121–1133
47 Chuang, C.F. and Meyerowitz, E.M. (2000) Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 97, 4985–4990
48 Kalidas, S. and Smith, D.P. (2002) Novel genomic cDNA hybrids produce effective RNA interference in adult *Drosophila*. *Neuron* 33, 177–184
49 Waterhouse, P.M. and Helliwell, C.A. (2003) Exploring plant genomes by RNA-induced gene silencing. *Nat. Rev. Genet.* 4, 29–38
50 Fuhrmann, M. *et al.* (2001) The abundant retinal protein of the *Chlamydomonas* eye is not the photoreceptor for phototaxis and photophobic responses. *J. Cell Sci.* 114, 3857–3863
51 Barth, C. and Jander, G. (2006) *Arabidopsis* myrosinases TGG1 and TGG2 have redundant function in glucosinolate breakdown and insect defense. *Plant J.* 46, 549–562
52 Robbins, T.P. *et al.* (1991) Meiotic instability of the *R-r* complex arising from displaced intragenic exchange and intrachromosomal rearrangement. *Genetics* 129, 271–283
53 Mourad, G. *et al.* (1994) Intragenic recombination in the *CSR1* locus of *Arabidopsis*. *Mol. Gen. Genet.* 243, 178–184
54 Ito, T. *et al.* (2005) A resource of 5,814 dissociation transposon-tagged and sequence-indexed lines of *Arabidopsis* transposed from start loci on chromosome 5. *Plant Cell Physiol.* 46, 1149–1153
55 Kuromori, T. *et al.* (2004) A collection of 11,800 single-copy *Ds* transposon insertion lines in *Arabidopsis*. *Plant J.* 37, 897–905
56 Eamens, A.L. *et al.* (2004) A bidirectional gene trap construct suitable for T-DNA and *Ds*-mediated insertional mutagenesis in rice (*Oryza sativa* L.). *Plant Biotechnol. J.* 2, 367–380
57 Sallaud, C. *et al.* (2004) High throughput T-DNA insertion mutagenesis in rice: a first step towards *in silico* reverse genetics. *Plant J.* 39, 450–464

58 Jeon, J.S. *et al.* (2000) T-DNA insertional mutagenesis for functional genomics in rice. *Plant J.* 22, 561–570
59 An, S. *et al.* (2003) Generation and analysis of end sequence database for T-DNA tagging lines in rice. *Plant Physiol.* 133, 2040–2047
60 Chen, S. *et al.* (2003) Distribution and characterization of over 1000 T-DNA tags in rice genome. *Plant J.* 36, 105–113
61 Tantikanjana, T. *et al.* (2004) Functional analysis of the tandem-duplicated P450 genes *SPS/BUS/CYP79F1* and *CYP79F2* in glucosinolate biosynthesis and plant development by *Ds* transposition-generated double mutants. *Plant Physiol.* 135, 840–848
62 Schaefer, D.G. and Zryd, J.P. (1997) Efficient gene targeting in the moss *Physcomitrella patens*. *Plant J.* 11, 1195–1206
63 Cove, D. (2005) The moss *Physcomitrella patens*. *Annu. Rev. Genet.* 39, 339–358
64 Beetham, P.R. *et al.* (1999) A tool for functional plant genomics: chimeric RNA/DNA oligonucleotides cause *in vivo* gene-specific mutations. *Proc. Natl. Acad. Sci. U. S. A.* 96, 8774–8778
65 Zhu, T. *et al.* (2000) Engineering herbicide-resistant maize using chimeric RNA/DNA oligonucleotides. *Nat. Biotechnol.* 18, 555–558
66 Zhu, T. *et al.* (1999) Targeted manipulation of maize genes *in vivo* using chimeric RNA/DNA oligonucleotides. *Proc. Natl. Acad. Sci. U. S. A.* 96, 8768–8773
67 Shaked, H. *et al.* (2005) High-frequency gene targeting in *Arabidopsis* plants expressing the yeast RAD54 gene. *Proc. Natl. Acad. Sci. U. S. A.* 102, 12265–12269
68 Eisen, J.A. *et al.* (1995) Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Res.* 23, 2715–2723
69 Kim, Y.G. *et al.* (1996) Hybrid restriction enzymes: zinc finger fusions to *Fok* I cleavage domain. *Proc. Natl. Acad. Sci. U. S. A.* 93, 1156–1160
70 Lloyd, A. *et al.* (2005) Targeted mutagenesis using zinc-finger nucleases in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* 102, 2232–2237
71 Wright, D.A. *et al.* (2005) High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant J.* 44, 693–705
72 Mandell, J.G. and Barbas, C.F. (2006) Zinc finger tools: custom DNA-binding domains for transcription factors and nucleases. *Nucleic Acids Res.* 34, 516–523
73 Dhanasekaran, M. *et al.* (2006) Designer zinc finger proteins: tools for creating artificial DNA-binding functional proteins. *Acc. Chem. Res.* 39, 45–52
74 Wright, D.A. *et al.* (2006) Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly. *Nat. Protoc.* 1, 1637–1652