The Plant Journal (2001) **27**(3), 235–242

TECHNICAL ADVANCE

A fast neutron deletion mutagenesis-based reverse genetics system for plants

Xin Li[1], Yujuan Song[1], Karen Century[1,†], Shelly Straight[1], Pamela Ronald[2], Xinnian Dong[3], Michael Lassner[4,‡] and Yuelin Zhang[1*]

[1]*Maxygen-Davis, 1105 Kennedy Place, Suite 5, Davis, CA 95616, USA,*
[2]*Department of Plant Pathology, University of California, Davis, CA 95616, USA,*
[3]*Developmental, Cell, and Molecular Biology Group, Department of Biology, Box 91000, Duke University, Durham, NC 27708–1000, USA, and*
[4]*Maxygen, Inc., 515 Galveston Drive, Redwood City, CA 94063, USA*

Received 3 April 2001; revised 9 May 2001; accepted 9 May 2001.
*For correspondence (fax +1 530 759 9896; e-mail yeulin.zhang@maxygen.com).
†Present address: DNA Plant Technology, 6701 San Pablo Avenue, Oakland, CA 94608, USA.
‡For obtaining deletion mutants (fax +650 3 64 2715; e-mail mike.lassner@maxygen.com)

Summary

A new reverse genetics method has been developed to identify and isolate deletion mutants for targeted plant genes. Deletion mutant libraries are generated using fast neutron bombardment. DNA samples extracted from the deletion libraries are used to screen for deletion mutants by polymerase chain reaction (PCR) using specific primers flanking the targeted genes. By adjusting PCR conditions to preferentially amplify the deletion alleles, deletion mutants were identified in pools of DNA samples, each pool containing DNA from 2592 mutant lines. Deletion mutants were obtained for 84% of targeted loci from an *Arabidopsis* population of 51 840 lines. Using a similar approach, a deletion mutant for a rice gene was identified. Thus we demonstrate that it is possible to apply this method to plant species other than *Arabidopsis*. As fast neutron mutagenesis is highly efficient, it is practical to develop deletion mutant populations with more complete coverage of the genome than obtained with methods based on insertional mutagenesis. Because fast neutron mutagenesis is applicable to all plant genetic systems, this method has the potential to enable reverse genetics for a wide range of plant species.

Keywords: reverse genetics, fast neutron mutagenesis, PCR, deletion mutants, genomics, plants.

Introduction

With the completion of the *Arabidopsis* genome sequencing effort (Arabidopsis Genome Initiative, 2000), the sequence of every *Arabidopsis* gene is now known. An international consortium is currently focusing on rice, and in the near future the sequence of the complete rice genome will also be available (Sasaki and Burr, 2000). DNA and protein sequence analyses have failed to identify the functions of the majority of *Arabidopsis* genes. The challenge for the post-sequencing era is to identify the biological functions of the sequenced genes in *Arabidopsis*, rice and other plant species. Reverse genetics will play an essential role in the process of assigning functions to a large number of unknown genes.

Gene silencing by antisense or sense suppression is a widely used method for probing the functions of plant genes (Baulcombe, 1996). Recently RNAi and intron-spliced hairpin have been shown to be quite effective in silencing endogenous genes (Chuang and Meyerowitz, 2000; Smith *et al.*, 2000). A problem associated with gene-silencing strategies is that the targeted gene is often only partially inactivated. As it is not possible to predict the extent of target gene disruption, data interpretation is difficult (Höfgen *et al.*, 1994; van der Krol *et al.*,

© 2001 Blackwell Science Ltd

1990). In addition, as transgenic plants need to be generated for gene silencing, large-scale characterization of genes with unknown functions requires creating a significant number of transgenic plants, which is impractical for crop species.

Insertion of T-DNA and transposable elements into coding or regulatory sequences of a gene can often disrupt gene function (Feldmann, 1991; Martienssen, 1998). Insertions within a targeted gene can be easily identified by polymerase chain reaction (PCR) using a combination of a gene-specific primer and a primer complementary to the T-DNA or transposon border sequences. A large number of insertion lines have been generated in *Arabidopsis* (Azpiroz-Leehan and Feldmann, 1997; Bouchez and Höfte, 1998; Koncz *et al*., 1992; Parinov *et al*., 1999; Speulman *et al*., 1999; Tissier *et al*., 1999; Wisman *et al*., 1998a; Wisman *et al*., 1998b), and reverse genetics screening was successfully applied to identify insertions in targeted genes (Krysan *et al*., 1996; McKinney *et al*., 1995). While screening insertion mutants is relatively straightforward, the number of insertion lines required to saturate the whole genome is very large (Krysan *et al*., 1999). It is not a trivial task to obtain hundreds of thousands of T-DNA lines by transformation or independent transposon lines by positive/negative selection for transposition (Sundaresan *et al*., 1995; Tissier *et al*., 1999). As a result, there are still a significant number of *Arabidopsis* genes for which an insertion cannot be found in the publicly available insertion mutant libraries (Meissner *et al*., 1999). It is even more difficult to obtain insertions for genes with smaller sizes (Krysan *et al*., 1996).

More recently, the TILLING (targeting induced local lesions in genomes) method was developed to identify ethyl methanesulfonate (EMS)-induced mutants (McCallum *et al*., 2000a; McCallum *et al*., 2000b). One advantage of TILLING is that it identifies a range of missense alleles in addition to knockouts. Single base-pair substitutions can be useful in studying gene function. Although EMS is a very effective mutagen, and saturating the genome with deleterious mutations can be achieved with a relatively small number of plants, only a few plants can be screened in each PCR reaction, which lengthens the screening procedure. A similar approach was used in *Caenorhabditis elegans* to identify chemical-induced deletion mutants (Jansen *et al*., 1997; Liu *et al*., 1999). Although in each PCR reaction over a thousand genomes can be screened to identify deletion mutants in *Caenorhabditis elegans*, more than one million genomes often need to be screened to obtain a deletion in a specific targeted gene, as the frequency of deletion mutation induced by chemicals is very low.

In plants, fast neutrons have been shown to be a very effective mutagen (Koornneef *et al*., 1982). As about 2500 lines treated with fast neutron at a dose of 60 Gy are required to inactivate a gene once on average (Koornneef *et al*., 1982) and the *Arabidopsis* genome contains about 25 000 genes (Arabidopsis Genome Initiative, 2000), it is estimated that about 10 genes are randomly deleted in each line. Molecular characterization of *Arabidopsis ga1-3* (Sun *et al*., 1992) and tomato *prf-3* (Salmeron *et al*., 1996) further demonstrated that fast neutron bombardment induces deletion mutations. Here we describe a new reverse genetics system based on fast neutron mutagenesis. As fast neutron-treated lines are easy to generate, we were able rapidly to assemble an *Arabidopsis* deletion library that allowed us to find deletion mutants for targeted genes at a frequency of about 80%. We also demonstrated that the same approach can be used in rice. The mechanism for *Arabidopsis* researchers to gain access to deletions from the described population will be posted on the *Arabidopsis* newsgroup (bionet.genome.arabidopsis, also accessible at http://www.bio.net/hypermail/arab-gen/).

Results

Strategy for high-throughput PCR screening of fast neutron library

To obtain deletion mutants for targeted genes, random deletion libraries were produced by fast neutron mutagenesis and then screened for specific deletion mutants by PCR. While it is relatively easy to generate a large number of fast neutron lines, PCR screening of these lines is not as straightforward as screening insertion libraries. To screen an insertion library by PCR, one gene-specific primer is used in combination with a primer specific to the insertion element in order to discriminate amplification of the insertion from wild-type DNA in pools of over a thousand lines. Screening for a deletion mutant by PCR requires that both primers are specific to the targeted locus. Such primers can amplify both the wild-type gene and the mutant gene. In order to detect a mutant in a mixture of a large number of wild-type lines, the PCR extension time is shortened so that amplification of the wild-type fragment is suppressed.

To test whether we could detect a mutant in a mixture of more than a thousand wild-type plants, we performed a reconstruction experiment using a known deletion mutant *ga1-3* (Sun *et al*., 1992). In this experiment, we mixed *ga1-3* DNA with different ratios of wild-type plant DNA, and used PCR to amplify the *ga1* locus. Figure 1 shows that even with 1000-fold excess of wild-type DNA, we can still readily see the deletion mutant band. The 6.4 kb wild-type fragment is not amplified under these PCR conditions, and does not interfere with the detection of *ga1-3*.

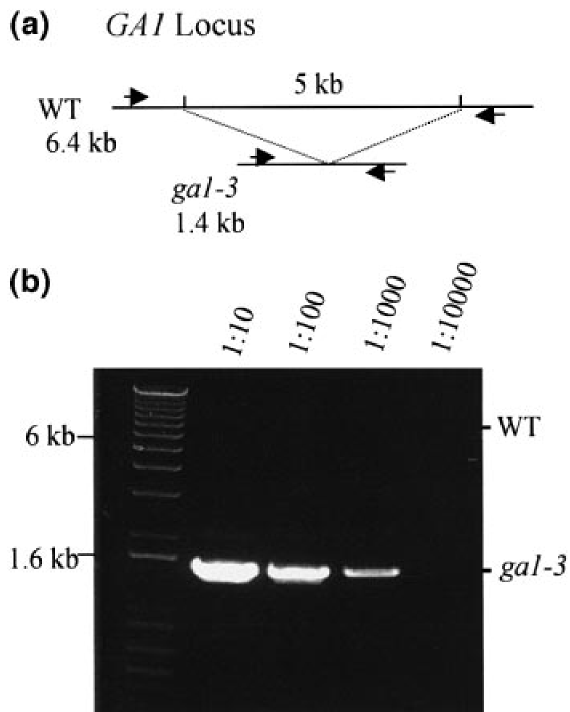
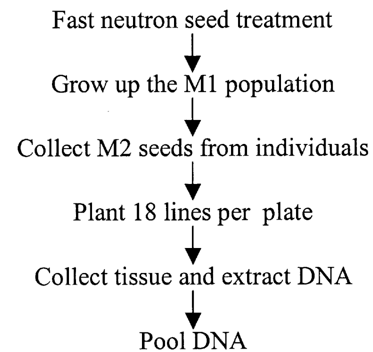


Figure 1. Reconstruction experiment using *Arabidopsis ga1-3* mutant DNA.
PCR was performed using a pair of primers flanking the *ga1-3* deletion. The distance between the two primers in the wild-type DNA is about 6.4 kb, and the extension time for the PCR reactions is 30 sec. Each reaction contains 0.5 units Takara *Ex* Taq polymerase, 100 ng wild-type DNA, and a different amount of *ga1-3* DNA as templates. The PCR products were analyzed by agarose gel electrophoresis followed by ethidium bromide staining. Lane 1, 1 kb plus ladder (Gibco-BRL); lanes 2–5, *ga1-3* DNA in the following amounts: lane 2, 10 ng; lane 3, 1 ng; lane 4, 0.1 ng; lane 5, 0.01 ng.



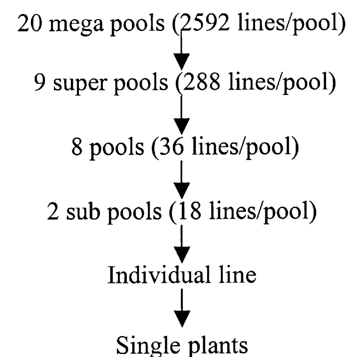


Figure 2. Deletion library construction and screening strategy.
(a) Flow chart for constructing a fast neutron deletion library.
(b) Flow chart of the PCR screening process showing the steps from initial analysis of the mega pools to finding single deletion mutant plants.

Deletion library construction and organization

To construct a deletion library (Figure 2a), wild-type *Arabidopsis* seeds were treated with fast neutrons. The treated M_1 seeds were planted, and M_2 seeds from individual plants were collected. About ten M_2 seeds were then taken from each line and the seeds from 18 lines were pooled together. These pooled seeds were plated on MS plates, and whole seedlings were collected after 8–9 days. Genomic DNA was isolated from the seedling tissue.

DNA samples representing all the mutant lines were aliquoted and organized into pools of increasing complexity (Figure 2b). Each mega pool contains DNA representing 2592 lines; each super pool contains DNA from 288 lines; and each pool contains DNA from 36 lines. The smaller pools allow rapid deconvolution once deletion mutants are found in any particular mega pool.

Finding a deletion mutant for AtMyb19

To identify a deletion mutant, a pair of primers specific to the sequence flanking the targeted gene and a pair of nested primers are selected using genomic DNA sequence information. The two pairs of primers are used in PCR to amplify the wild-type DNA fragment using a long extension time to check primer quality, and a short extension time to determine conditions that suppress the amplification of the wild-type fragment. The short extension time is then used to screen the mega pools. After the first round of PCR, nested PCR is performed on the 1 : 50 diluted products of the first-round PCR to increase the sensitivity and specificity. The product of the second-round PCR is checked using agarose gel electrophoresis to detect the presence of amplified DNA fragments derived from deletion alleles. If a deletion band is found in a mega pool, PCR analysis is subsequently carried out on the constituent super pools and pools (Figure 2b). Once a deletion is identified in a pool of 18 lines, seeds from those 18 lines are then planted individually, and DNA samples from the plants are analyzed by PCR to look for the single line that carries the deletion.

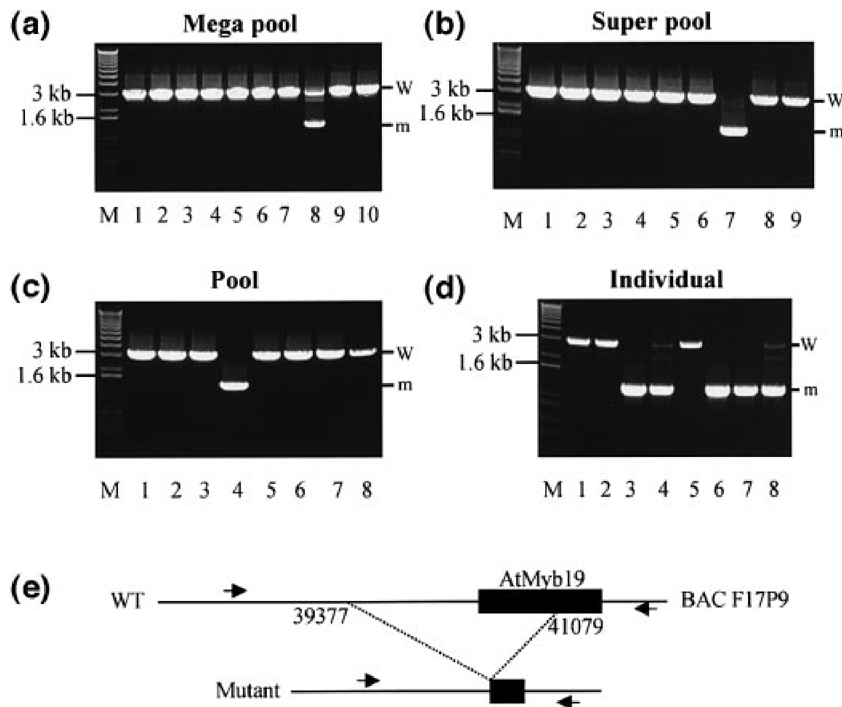


Figure 3. Identification of a deletion mutant for *AtMyb19*.
(a) PCR on DNA from the ten mega pools of 2592 lines is shown (1–10).
(b) PCR was performed on DNA from the nine super pools (1–9) of 288 lines that make up mega pool number 8.
(c) PCR was performed on DNA from the eight pools (1–8) of 36 lines that make up super pool number 7.
(d) Eight single plants (1–8) from the line containing the deletion mutant were analyzed by PCR. All PCR was performed using Qiagen HotStarTaq™ polymerase. Two rounds of PCR were performed on the DNA pools, and the nested PCR products were analyzed by agarose gel electrophoresis and stained with ethidium bromide. W, wild type; m, mutant. Arrows represent the location of the PCR primers.

The data presented in Figure 3 illustrate the isolation of a deletion mutant for *AtMyb19* (Romero *et al.*, 1998). *Myb19* is a transcription factor; a previous attempt to isolate mutants using insertional mutation failed (Meissner *et al.*, 1999). Ten mega pools, a population containing 25 920 lines, were screened as described above, and a deletion band was detected in mega pool number 8 (Figure 3a). DNA sequence analysis of the deletion PCR product revealed the deletion of a 1.7 kb fragment (corresponding to nucleotide 39377–41079 of BAC F17P19, accession number AB025603), including most of the *Myb19* coding sequences (Figure 3e).

PCR was performed on the nine constituent super pools of mega pool number 8. As shown in Figure 3(b), super pool number 7 was found to contain the deletion mutant. PCR analysis of the eight pools composing super pool 7 identified a pool of 36 lines containing the mutant (Figure 3c). Subsequently, a single line carrying the deletion was identified. In Figure 3(d), PCR was carried out on 8 M_2 plants from the single line that contains the deletion mutation. Both wild-type and mutant bands were amplified from plants 4 and 8. Only the deletion band was amplified from plants 3, 6 and 7; and only the wild-type band was amplified from plants 1, 2 and 5. These data indicate that plants 4 and 8 are heterozygous for the deletion allele; plants 3, 6 and 7 are homozygous for the deletion allele; and plants 1, 2 and 5 are homozygous for the wild-type allele.

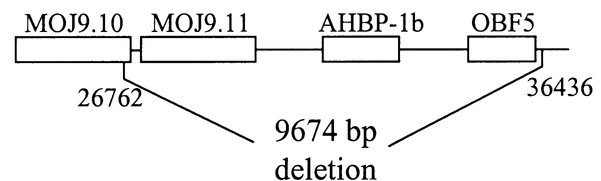


Figure 4. Structure of the deletion that knocks out two tandem bZIP transcription factors.
The junction of the deletion was determined by sequencing. A fragment of about 9.7 kb between nucleotides 26762 and 36436 (P1 clone MOJ9, AB010697) was deleted. The gene structures are drawn based on the annotation of clone MOJ9.

Deletion of two tandem transcription factors

Arabidopsis bZIP transcription factor AHBP-1b and OBF5 were previously shown to interact with NPR1 (Depres *et al.*, 2000; Zhang *et al.*, 1999; Zhou *et al.*, 2000), a key regulator of systemic acquired resistance. In the yeast two-hybrid assay, AHBP-1b binds to NPR1 with greater affinity than does OBF5. *AHBP-1b* and *OBF5* are directly linked on chromosome 5 (Figure 4). The distance between the coding regions of these two genes is less than 2 kb. To identify a plant with both genes deleted, we designed primers flanking the region containing these two genes. The distance between the primers is about 17 kb. Twenty mega pools representing 51 840 fast neutron lines were screened, and a deletion encompassing both genes was identified in a single mega pool. Individual plants contain-

Table 1. Size distribution of 36 *Arabidopsis* deletion alleles

Deletion size (kb)	Number of deletions
0–2	7
2–4	14
4–6	6
6–8	4
8–10	4
10–12	1

ing the tandem gene deletion were identified using the process described above. DNA sequence analysis showed that a fragment of about 9.7 kb, including both *AHBP-1b* and *OBF5*, was deleted in the mutant. In addition to the two transcription factors, the complete coding region of a putative receptor kinase and the C-terminal 20 amino acids of a hypothetical protein were also deleted. The location of the deletion is shown in Figure 4. Homozygous plants for this deletion were recovered and inoculated with *Pseudomonas syringae* pv. *tomato* DC3000 to test for enhanced susceptibility. No difference was observed between the mutant and the wild-type plants (data not shown). There are two closely related bZIP transcription factors, TGA3 and TGA6, that can also bind to NPR1 (Depres *et al.*, 2000; Zhang *et al.*, 1999; Zhou *et al.*, 2000). The *AHBP-1b* and *OBF5* deletion mutant may have to be combined with *TGA3* and *TGA6* mutants in order to observe altered resistance responses to pathogens.

Characterization of the Arabidopsis *deletion library*

To test the general applicability of the method, we screened the *Arabidopsis* population for deletions in 23 additional loci. One of these loci is a three-gene tandem array, and the rest are single gene targets. The sizes of the single genes (from start to stop codon) range from about 1 to 6 kb. Among the 23 single genes, eight are between 1 and 2 kb in length, nine are between 2 and 3 kb in length, and six are between 3 and 6 kb in length.

We identified deletion mutants for 21 of the 25 loci, including the three-gene tandem array. For each of these 21 loci we found at least one mutation that deletes part or all of the coding sequences. The deletions range in size from 0.8 to 12 kb, as detailed in Table 1. We identified two deletion alleles for nine of the loci, three deletion alleles for three other loci, and one deletion allele for the remaining nine loci.

Construction of a rice deletion library and identification of a deletion mutant for a rice gene

To test the applicability of the deletion-based reverse genetics system in rice, a preliminary population consisting of 24 660 fast neutron lines was generated. DNA samples were isolated from the M_2 plants, and organized into pools similar to those described for the *Arabidopsis* population. We screened for deletions in five targeted genes using the methodology described above, and identified a deletion mutant for one of the targets (*RG1*, accession number BAA89552) which has no known function. As shown in Figure 5(a), the deletion for *RG1* was first identified in mega pool number 10. Mega pools 1–9 contain DNA from 2592 lines, while mega pool 10 contains DNA from the other 1332 lines. Screening the super pools composing mega pool 10 identified a single super pool of

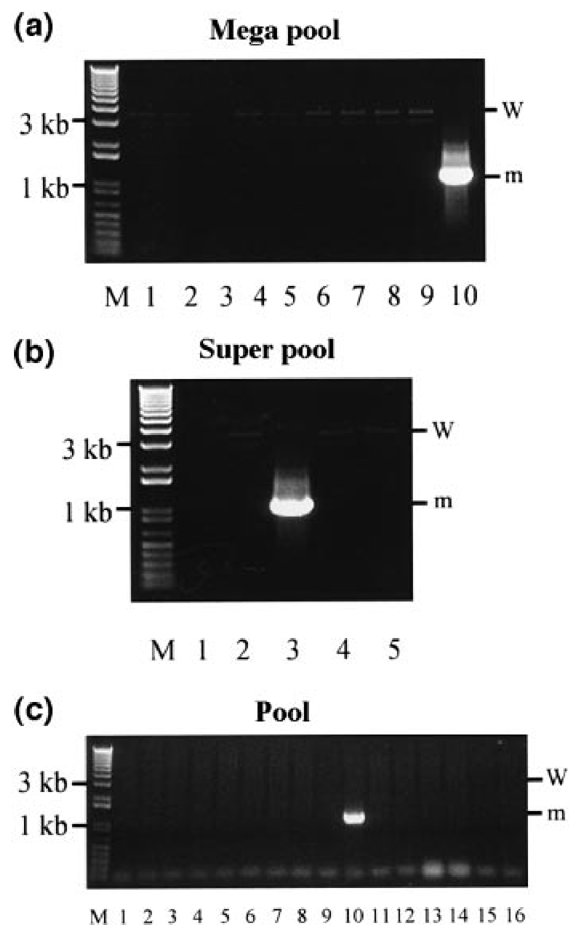


Figure 5. Identification of a rice deletion mutant.
(a) PCR on DNA of rice mega pools (1–10) is shown. Each mega pool contains 2592 rice fast neutron lines, except that mega pool 10 has only 1332 lines.
(b) PCR was performed on DNA of five super pools (1–5) that make up mega pool number 10.
(c) PCR was performed on DNA of 16 pools of 18 lines (1–16) that make up super pool number 3. Two rounds of PCR were performed on the DNA of mega pools and super pools. The nested PCR products were loaded on agarose gels and stained with ethidium bromide after electrophoresis. Only one round of PCR was performed on DNA of pools of 18 plants. All PCR was performed using Qiagen HotStarTaq™ polymerase and 1 min 30 sec extension time. W, wild type; m, mutant.

288 lines carrying the deletion (Figure 5b). PCR analysis on the 16 constituent sub-pools identified a pool of 18 lines containing the deletion mutant (Figure 5c). Sequence analysis of the mutant DNA band showed that a fragment of about 2.5 kb was deleted in the gene.

Discussion

We have demonstrated that deletion mutants can be identified for targeted plant genes by screening fast neutron-mutagenized populations via an efficient PCR screening procedure. A key advantage of the deletion-based reverse genetics system over insertional mutagenesis-based methods is that fast neutron mutagenesis can be performed on a large number of dry seeds, and that no plant transformation is required. During the past few years, different groups have generated large numbers of T-DNA and transposon insertion lines (Azpiroz-Leehan and Feldmann, 1997; Bouchez and Höfte, 1998; Koncz *et al.*, 1992; Parinov *et al.*, 1999; Speulman *et al.*, 1999; Tissier *et al.*, 1999; Wisman *et al.*, 1998a; Wisman *et al.*, 1998b). If all these lines are screened, knockout plants can probably be found for most *Arabidopsis* genes. However, saturation of the *Arabidopsis* genome with insertion elements is far from complete, and screening all the available lines represents a logistical challenge. To have 99% probability of finding an insertion in a 1 kb gene, about 550 000 insertion lines would need to be screened (Krysan *et al.*, 1999).

In a population of 51 840 fast neutron lines, we found deletion mutants for more than 80% of the 25 loci tested. Based on these data, we estimate that a population size of 84 825 will enable a success rate of 95% in isolating deletions in target genes, and a population of 130 397 will yield a 99% probability of success in isolating a deletion in any target locus (based on the formula $N = \ln[1-P]/\ln[1-F]$, where N is the population size, P is the probability of isolating a deletion, and F is the frequency of deletions that can be isolated using the deletion-based reverse genetics system; F was calculated using the data presented in this manuscript, $N = 51\,840$, and $P = 0.84$).

A critical factor in finding a knockout plant in a large mutant population is the efficiency of the screening process. Higher throughput often means an increased number of lines can be screened with a fixed cost. The efficiency of the PCR screening process for deletions is comparable to that of the screening process for insertion lines. As fewer fast neutron lines need to be screened in order to find a mutant for a targeted gene, it may actually take less time to screen for a deletion than to screen for an insertion. The identification of mutants using TILLING (McCallum *et al.*, 2000a; McCallum *et al.*, 2000b) requires smaller population sizes than are required for the method described here. However, as TILLING can screen only a small number of lines in each PCR reaction, the overall efficiency of the deletion method is higher.

To reach the defined goal of obtaining mutants for every *Arabidopsis* gene and understanding their functions by 2010 (Somerville and Dangl, 2000), different and complementary approaches are necessary. Characterization of the insertion sites of transposons (Parinov *et al.*, 1999; Tissier *et al.*, 1999) and T-DNAs by DNA sequence analysis can be carried out in higher throughput than screening mutant populations by any kind of pooling strategy. It is anticipated that plants harboring mutations in many genes will be identified by searching a DNA sequence database with a collection of a large number of insertion site sequences once such a database is established. However, insertion lines for a significant percentage of genes may not be available using this approach.

The probability of finding transposon or T-DNA insertion mutants is directly proportional to the size of the target gene (Krysan *et al.*, 1999). Thus it will be a challenge to find insertions in genes smaller than 1 kb. As deletion mutations that can be detected by PCR screening often affect larger regions than insertion mutations, it will be easier to hit a small gene with a deletion than with an insertion. In our screen for 25 target loci, we were able to find deletions for genes of different sizes. None of the four genes we failed to obtain deletion mutants for was smaller than 2 kb. This indicates that small gene size is not the reason for our failure to detect deletions for those four genes. We believe that fast neutron mutagenesis will be very useful for isolating lines harboring mutations in small genes.

Another challenge for reverse genetics is genes in tandem arrays. The *Arabidopsis* genome has 1528 tandem arrays containing 4140 individual genes (Arabidopsis Genome Initiative, 2000). Potential functional redundancy encoded by related genes in these arrays can be a problem for functional characterization of *Arabidopsis* genes by reverse genetics. We have demonstrated that it is possible to find deletions mutating two or three tandem homologous genes.

One potential limitation for using deletion mutants obtained with fast neutron mutagenesis is that some deletions can affect more than one gene. For example, the 9.7 kb deletion mutant that knocks out *AHBP-1b* and *OBF5* also removes a putative receptor kinase gene. When a phenotype is observed in a mutant that deletes more than one gene, complementation tests will need to be performed to determine mutations in which gene is responsible for the phenotype. In some cases this can also be resolved by analyzing another independent mutant allele. As gene density in *Arabidopsis* (about one gene every 4.8 kb) is quite high (Bevan *et al.*, 1998), deletions larger than 3 kb have a high probability of knocking out more than one gene. The problem associated with deletions affecting multiple genes is less critical in rice, as both the

gene sizes and the intergenic regions in rice are generally larger than in *Arabidopsis*. Another potential problem associated with large deletions is poor transmission through male gametes. The size limit of transmittable deletions is unclear. For the 36 deletions (<13 kb) we obtained, this does not appear to be a problem.

In principle, the deletion-based reverse genetics system can be applied to most plant species. Unlike insertional mutagenesis or gene silencing, the deletion-based method requires genomic sequence information for a relatively large region around the gene of interest. The lack of genomic sequences around a target gene can be a time-limiting factor in crop plants. An international consortium is currently determining the DNA sequence of the rice genome. This will enable large-scale reverse genetics in rice. As rice has a genome size at least three times larger than *Arabidopsis*, and transformation in rice is not nearly as efficient as in *Arabidopsis*, it will be very difficult to saturate the rice genome with T-DNA or transposable elements. We have demonstrated the application of the fast neutron-based reverse genetics system to rice. It is anticipated that we can cover the whole rice genome with easily detectable deletion mutations simply by expanding the rice fast neutron population.

In crop plants, this method can also potentially be used to inactivate unwanted genes in order to generate desirable phenotypes for agriculture. As genetic transformation is not used in the process, the products will contain no foreign DNA. Thus the resultant crop varieties will not face the regulatory or public acceptance barriers associated with transgenic crops.

Experimental procedures

Plant materials

Wild-type Columbia *Arabidopsis* seeds were treated by fast neutron bombardment at a dose of 60 Gy by Andrea Kodym (International Atomic Energy Agency Agriculture and Biotechnology Laboratory, Vienna, Austria) and Joe Palfalvi (Atomic Energy Research Institute, Budapest, Hungary). M_1 plants were grown on soil and allowed to self-pollinate. Seeds from individual plants were collected. The plants were grown at 22°C under 16 h light/8 h dark cycles. For M_2 mutant seedlings, sterilized seeds were plated on Murashige–Skoog (MS) medium (Murashige and Skoog, 1962). After 8–9 days, whole seedlings were harvested for DNA extraction. To mutagenize rice seeds, wild-type *M*202 seeds were treated by fast neutron bombardment at different doses ranging from 18–30 Gy. The rice seedlings were planted in a greenhouse and manually transplanted to the field 5 weeks after germination. Seeds from individual plants were harvested.

Genomic DNA purification

Arabidopsis genomic DNA was extracted using a CTAB-based protocol (Dellaporta *et al.*, 1983), and rice genomic DNA was extracted using a slightly modified procedure (Chen and Ronald, 1999). Aliquots of DNA samples from 36 *Arabidopsis* lines or 18 rice lines were pooled and further purified using a DNeasy plant Mini Kit (Qiagen Inc., Valencia, CA, USA).

PCR and sequencing analysis

PCR was performed using either HotStarTaq DNA polymerase (Qiagen) for short fragments (<6 kb) or Takara *Ex* Taq polymerase (Panvera Corp., Madison, WI, USA) for large fragments (>6 kb). The distances between primer pairs used in the screening range from 3 to 17 kb, depending on the targeted loci. For routine attempts to achieve deletion in a single gene, screenings are done using primers either 3b or 9 kb apart. The extension time required for amplification of the wild-type DNA was first checked empirically for each pair of primers, then screening for deletion mutants was performed using an extension time that suppressed the amplification of the wild-type fragments. For example, screening with primers 3 kb apart was carried out using a 90 sec extension time, and screening with primers 17 kb apart was carried out using a 3 min extension time. PCR products were analyzed on 1% agarose gels and deletion mutant DNA bands were purified using QIAquick Gel Extraction Kit (Qiagen). The products were sequenced using an ABI 377 automated sequencer, and sequences were compared to wild-type sequences to find the deletion junctions.

Acknowledgements

We thank Kelly Lin, Ryan Blenkush, Kaunghtet Myat and Joanne Greenberg for assistance in daily laboratory work. David Mackill's and Pam Ronald's labs are thanked for their help with the rice field work. Walker Lutringer and Troy Obrero are thanked for DNA sequencing. We greatly appreciate Peter Quail, John Bedbrook and David McElroy for critical reading and helpful comments on the manuscript.

References

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

Azpiroz-Leehan, R. and Feldmann, K.A. (1997) T-DNA insertion mutagenesis in *Arabidopsis*: going back and forth. *Trends Genet.* **13**, 152–156.

Baulcombe, D.C. (1996) RNA as a target and an initiator of post-transcriptional gene silencing in transgenic plants. *Plant Mol. Biol.* **32**, 79–88.

Bevan, M., Bancroft, I., Bent, E. *et al.* (1998) Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. *Nature*, **391**, 485–488.

Bouchez, D. and Höfte, H. (1998) Functional genomics in plants. *Plant Physiol.* **118**, 725–732.

Chen, D. and Ronald, P.C. (1999) A rapid DNA minipreparation method suitable for AFLP and other PCR applications. *Plant Mol. Biol. Rep.* **17**, 53–57.

Chuang, C.-F. and Meyerowitz, E.M. (2000) Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA*, **97**, 4985–4990.

Dellaporta, S.L., Wood, J. and Hicks, J.B. (1983) A plant DNA minipreparation, Version II. *Plant Mol. Biol. Rep.* **1**, 19–21.

Depres, C., DeLong, C., Glaze, S., Liu, E. and Fobert, P.R. (2000) The *Arabidopsis* NPR1/NIM1 protein enhances the DNA binding activity of a subgroup of the TGA family of bZIP transcription factors. *Plant Cell*, **12**, 279–290.

Feldmann, K.A. (1991) T-DNA insertion mutagenesis in *Arabidopsis*: mutational spectrum. *Plant J.* **1**, 71–82.

Höfgen, R., Axelsen, K.B., Kannangara, C.G., Schuttke, I., Pohlenz, H.D., Willmitzer, L., Grimm, B. and Wettstein, D.V. (1994) A visible marker for antisense mRNA expression in plants: inhibition of chlorophyll synthesis with a glutamate-1-semialdehyde aminotransferase antisense gene. *Proc. Natl Acad. Sci. USA*, **91**, 1726–1730.

Jansen, G., Hazendonk, E., Thijssen, K.L. and Plasterk, R.H.A. (1997) Reverse genetics by chemical mutagenesis in *Caenorhabditis elegans*. *Nature Genet.* **17**, 119–121.

Koncz, C., Nemeth, K., Redei, G.P. and Schell, J. (1992) T-DNA insertional mutagenesis in *Arabidopsis*. *Plant Mol. Biol.* **20**, 963–976.

Koornneef, M., Dellaert, L.W.M. and van der Veen, J.H. (1982) EMS- and radiation-induced mutation frequencies at individual loci in *Arabidopsis thaliana* (L.) Heynh. *Mutat. Res.* **93**, 109–123.

van der Krol, A.R., de Mur, L.A., Lange, P., Gerats, A.G.M., Mol, J.N.M. and Stuitje, A.R. (1990) Antisense chalcone synthase genes in petunia: visualization of variable transgene expression. *Mol. Gen. Genet.* **220**, 204–212.

Krysan, P.J., Young, J.C., Tax, F. and Sussman, M.R. (1996) Identification of transferred DNA insertions within *Arabidopsis* genes involved in signal transduction and ion transport. *Proc. Natl Acad. Sci. USA*, **93**, 8145–8150.

Krysan, P.J., Young, J.C. and Sussman, M.R. (1999) T-DNA as an insertional mutagen in *Arabidopsis*. *Plant Cell*, **11**, 2283–2290.

Liu, L.X., Spoerke, J.M., Mulligan, E.L. *et al.* (1999) High-throughput isolation of *Caenorhabditis elegans* deletion mutants. *Genome Res.* **9**, 859–867.

Martienssen, R.A. (1998) Functional genomics: probing plant gene function and expression with transposons. *Proc. Natl Acad. Sci. USA*, **95**, 2021–2026.

McCallum, C.M., Comai, L., Greene, E.A. and Henikoff, S. (2000a) Targeted screening for induced mutations. *Nature Biotechnol.* **18**, 455–457.

McCallum, C.M., Comai, L., Greene, E.A. and Henikoff, S. (2000b) Targeting induced local lesions in genomes (TILLING) for plant functional genomics. *Plant Physiol.* **123**, 439–442.

McKinney, E.C., Ali, N., Traut, A., Feldmann, K.A., Belostotsky, D.A., McDowell, J.M. and Meagher, R.B. (1995) Sequence-based identification of T-DNA insertion mutations in *Arabidopsis*: actin mutants *act2-1* and *act4-1*. *Plant J.* **8**, 613–622.

Meissner, R.C., Jin, H., Cominelli, E. *et al.* (1999) Function search in a large transcription factor gene family in *Arabidopsis*: assessing the potential of reverse genetics to identify insertional mutations in R2R3 *MYB* genes. *Plant Cell*, **11**, 1827–1840.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant*, **15**, 473–497.

Parinov, S., Sevugan, M.Y.e.D., Yang, W.-C., Kumaran, M. and Sundaresan, V. (1999) Analysis of flanking sequences from *Dissociation* insertion lines: a database for reverse genetics in *Arabidopsis*. *Plant Cell*, **11**, 2263–2270.

Romero, I., Fuertes, A., Benito, M.J., Malpica, J.M., Leyva, A. and Paz-Ares, J. (1998) More than *80R2R3-MYB* regulatory genes in the genome of *Arabidopsis thaliana*. *Plant J.* **14**, 273–284.

Salmeron, J.M., Oldroyd, G.E.D., Rommens, C.M.T., Scofield, S.R., Kim, H., Lavelle, D.T., Dahlbeck, D. and Staskawicz, B.J. (1996) Tomato *Prf* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. *Cell*, **86**, 123–133.

Sasaki, T. and Burr, B. (2000) International Rice Genome Sequencing Project: the effort to completely sequence the rice genome. *Curr. Opin. Plant Biol.* **3**, 138–141.

Smith, N.A., Singh, S.P., Wang, M.B., Stoutjesdijk, P.A., Green, A.G. and Waterhouse, P.M. (2000) Total silencing by intron-spliced hairpin RNAs. *Nature*, **407**, 319–320.

Somerville, C. and Dangl, J. (2000) Plant biology in 2010. *Science*, **290**, 2077–2078.

Speulman, E., Metz, P.L.J., van Arkel, G., te Lintel Hekkert, B., Stiekema, W.J. and Pereira, A. (1999) A two-component *Enhancer–Inhibitor* transposon mutagenesis system for functional analysis of the *Arabidopsis* genome. *Plant Cell*, **11**, 1853–1866.

Sun, T.P., Goodman, H.M. and Ausubel, F.M. (1992) Cloning the *Arabidopsis GA1* locus by genomic subtraction. *Plant Cell*, **4**, 119–128.

Sundaresan, V., Springer, P., Volpe, T., Haward, S., Jones, J.D., Dean, C., Ma, H. and Martienssen, R. (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. *Genes Dev.* **9**, 1797–1810.

Tissier, A.F., Marillonnet, S., Klimyuk, V., Patel, K., Torres, M.A., Murphy, G. and Jones, J.D.G. (1999) Multiple independent defective *Suppressor–mutator* transposon insertions in *Arabidopsis*: a tool for functional genomics. *Plant Cell*, **11**, 1841–1852.

Wisman, E., Cardon, G.H., Fransz, P. and Saedler, H. (1998a) The behaviour of the autonomous maize transposable element *En/Spm* in *Arabidopsis thaliana* allows efficient mutagenesis. *Plant Mol. Biol.* **37**, 989–999.

Wisman, E., Hartmann, U., Sagasser, M., Baumann, E., Palme, K., Hahlbrock, K., Saedler, H. and Weisshaar, B. (1998b) Knock-out mutants from an *En-1* mutagenized *Arabidopsis thaliana* population generate phenylpropanoid biosynthesis phenotypes. *Proc. Natl Acad. Sci. USA*, **95**, 12432–12437.

Zhang, Y., Fan, W., Kinkema, M., Li, X. and Dong, X. (1999) Interaction of NPR1 with basic leucine zipper protein transcription factors that bind sequences required for salicylic acid induction of the *PR-1* gene. *Proc. Natl Acad. Sci. USA*, **96**, 6523–6528.

Zhou, J.-M., Trifa, Y., Silva, H., Pontier, D., Lam, E., Shah, J. and Klessig, D.F. (2000) NPR1 differentially interacts with members of the TGA/OBF family of transcription factors that bind an element of the *PR-1* gene required for induction by salicylic acid. *Mol. Plant–Microbe Interact.* **13**, 191–202.