

Reverse genetic approaches for functional genomics of rice

Gynheung An*, Dong-Hoon Jeong, Ki-Hong Jung and Sichul Lee

National Research Laboratory of Plant Functional Genomics, Division of Molecular and Life Sciences, Pohang University of Science and Technology (POSTECH), Pohang 790-784, Korea (*author for correspondence; e-mail genean@postech.ac.kr)

Received 13 April 2004; accepted in revised form 30 September 2004

Key words: functional genomics, gene tagging, reverse genetics, transposon, T-DNA

Abstract

T-DNA and transposable elements, e.g., *Ds* and *Tos17*, are used to generate a large number of insertional mutant lines in rice. Some carry the *GUS* or *GFP* reporter for gene trap or enhancer trap. These reporter systems are valuable for identifying tissue- or organ-preferential genes. Activation tagging lines have also been generated for screening mutants and isolating mutagenized genes. To utilize these resources more efficiently, tagged lines have been produced for reverse genetic approaches. DNA pools of the T-DNA tagged lines and *Tos17* lines have been prepared for PCR screening of insertional mutants in a given gene. Tag end sequences (TES) of the inserts have also been produced. TES databases are beneficial for analyzing the function of a large number of rice genes.

Abbreviations: iPCR, inverse PCR; RT-PCR, reverse transcription-PCR; TAIL PCR, thermal asymmetric interlaced PCR; T-DNA, transferred DNA; UTR, untranslated region

Introduction

Sequencing of the rice genome is nearly complete (Goff *et al.*, 2002; Sasaki *et al.*, 2002; Yu *et al.*, 2002). Therefore, one of the most challenging goals now is to examine the functioning of its large number of genes. To facilitate such an evaluation, several reverse genetic approaches have been developed, including homologous recombination (reviewed in Hanin and Paszkowski, 2003), anti-sense or RNAi suppression (Chuang and Meyerowitz, 2000), and insertional mutagenesis (Feldmann, 1991; Jeon *et al.*, 2000). Among these, random insertional mutagenesis by transposons or T-DNA has been most widely used for large-scale analyses. This technique is not only efficient for identifying knockout mutants, but can also be employed for promoter trapping and activation tagging. The recent establishment of a large number of insertional mutants will accelerate these

reverse genetics approaches for studying gene function in this model monocot species.

T-DNA tagging

In *Arabidopsis*, T-DNA has been broadly utilized for generating insertional mutant pools and studying gene function (Azpiroz-Leehan and Feldmann, 1997; Bouche and Bouchez, 2001). For example, the *Arabidopsis* knockout facility at the University of Wisconsin has established a population of 60 480 T-DNA-tagged lines (Krysan *et al.*, 1999). Recently, Over 225 000 independent T-DNA insertional lines of *Arabidopsis* have been created that represent almost the entire gene space (Alonso *et al.*, 2003). T-DNA insertion into plant chromosomes is not biased toward a particular gene group, thereby allowing near saturation mutagenesis of the genome. Because *Arabidopsis*

can be easily transformed for the generation of a large quantity of transgenic plants, it is not too difficult to establish a reasonable number of knockout lines. However, applying this strategy to other species has been complicated because transformation efficiency is not high enough to produce the many thousand requisite transgenic plants.

An efficient means for rice transformation is via the *Agrobacterium*-mediated cocultivation method (Hiei *et al.*, 1994; Lee *et al.*, 1999). This technique allows researchers to obtain a suitable T-DNA tagging population. For example, An's group has generated approximately 100 000 fertile rice lines tagged by T-DNA (Jeon *et al.*, 2000; Jeong *et al.*, 2002), while work by Zhang and colleagues has resulted in more than 30 000 T-DNA insertional lines (Wu *et al.*, 2003). Several other groups have independently produced T-DNA insertional mutant lines in this species (Yin and Wang, 2000; Chen *et al.*, 2003; Sallaud *et al.*, 2003; Sha *et al.*, 2004). Because the average copy number of T-DNA inserts per line is 1.4–2.0, more than 300 000 T-DNA tags have now been generated in rice (Table 1). These populations are large enough to find a knockout in a given gene at more than 90% probability, assuming that T-DNA is randomly inserted into a chromosome. Nevertheless, the number of tagging lines necessary for saturating all the rice genes may be smaller than the estimated value because T-DNA insertions have been found to occur preferentially in gene-rich regions (Barakat *et al.*, 2000; An *et al.*, 2003; Chen *et al.*, 2003; Wu *et al.*, 2003).

Transposon tagging

Transposon tagging has been used to isolate numerous genes (Sundaresan, 1996) since the first successful cloning of the *bronze* locus in maize by the *Ac/Ds* (*Activator-Dissociation*) transposon system (Fedoroff *et al.*, 1983). Other systems, such as *En/Spm* (*Enhancer/Suppressor-mutator*) and *Mu* (*Mutator*), have also been adopted for cloning several maize genes (Walbot, 1992, 2000).

The maize *Ac/Ds* system has also been utilized for gene tagging in rice. Originally, the autonomous *Ac* element was cloned between a promoter and the *hygromycin phosphotransferase* coding region. The construct was then introduced into rice chromosomes by direct transformation. Transposition of the *Ac* element was proven by the recovery of hygromycin-resistant plants (Izawa *et al.*, 1991; Murai *et al.*, 1991) and newly transposed *Ac* insertions (Enoki *et al.*, 1999). Molecular analysis of this *Ac* transmission to the progeny revealed that their germinal transmission was detected in some insertions but not in others, indicating that some *Ac* does not transmit to the progeny.

The non-autonomous *Ds* element was also transposed in the presence of *Ac* transposase (Shimamoto *et al.*, 1993; Sugimoto *et al.*, 1994). Germinal transposition of *Ds* was observed at high frequency when a transgenic plant containing the *Ds* element was crossed with a transgenic plant constitutively expressing *Ac* transposase (Figure 1A). The frequency of *Ds* transposition

Table 1. Inventory of insertional mutants in rice.

Insertional element	Character	Number of lines	Web site	Reference
Ds	Gene trap	30 000	http://bk21.gsnu.ac.kr/cdhan	Chin <i>et al.</i> (1999)
Ds	Knock out	8 000	http://www.pi.csiro.au/fgrttpub/	Upadhyaya <i>et al.</i> (2002)
Ds	Knock out	18 000	http://www.tll.org.sg/res/sri.asp	Kolesnik <i>et al.</i> (2004)
Ds	Enhancer trap	10 000	http://genoplante-info.infobiogen.fr/oryzatagline/project.htm	Greco <i>et al.</i> (2003)
T-DNA	Gene trap	50 000	http://www.postech.ac.kr/life/pfg	Jeon <i>et al.</i> (2000)
T-DNA	Activation/ gene trap	50 000	http://www.postech.ac.kr/life/pfg	Jeong <i>et al.</i> (2002)
T-DNA	Enhancer trap	46 000	http://genoplante-info.infobiogen.fr/oryzatagline/project.htm	Sallaud <i>et al.</i> (2003)
T-DNA	Enhancer trap	42 000	http://www.genomics.zju.edu.cn/ricetdna	Chen <i>et al.</i> (2003)
Tos17	Knock out	50 000	http://tos.nais.affrc.go.jp	Miyao <i>et al.</i> (2003)

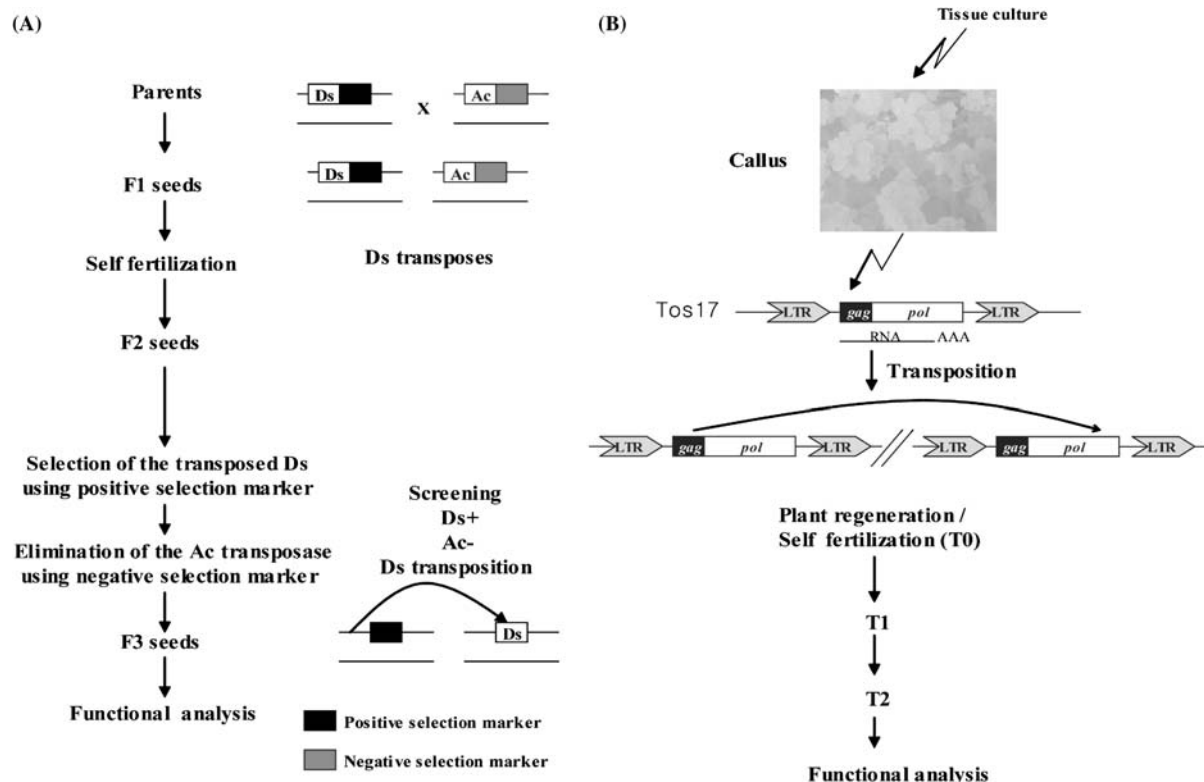


Figure 1. Schematic diagram of transposon tagging by AC/DS (A) and Tos17 (B).

significantly declined in subsequent generations, even in the presence of *Ac* transposase. However, transposition could be recovered in certain lines through protoplast regeneration (Kim *et al.*, 2002).

As shown in Table 1, *Ac/Ds* systems have been used for large-scale generation of gene tagging via the *Agrobacterium*-mediated gene delivery method (Chin *et al.*, 1999; Nakagawa *et al.*, 2000; Upadhyaya *et al.*, 2002; Kolesnik *et al.*, 2004). The germinal transposition frequency of *Ds* is high. Study of the transposition pattern in siblings has revealed that 79% contain at least two different insertions, suggesting late transposition during rice development (Kolesnik *et al.*, 2004). Moreover, repetitive ratoon culturing causes new transposition, at about 30% frequency, demonstrating that this strategy can produce a large mutant population (Chin *et al.*, 1999).

Retrotransposon tagging

Retroelements are a major component of plant genomes. At least 17% of the rice genome consists

of retroposons (McCarthy *et al.*, 2002). These are believed to be involved in gene duplication as well as the regulation of gene expression in both plants and animals (Nekrutenko and Li, 2001; Jordan *et al.*, 2003; Kashkush *et al.*, 2003). Transposition of some retroelements can be induced by stresses caused by pathogen infection, cell culture, and wounding (Hirochika *et al.*, 1996; Takeda *et al.*, 1998).

One such element in rice is *Tos17* (Figure 1B). Although its copy number is low (one to five, depending on the cultivar), this can increase to up to 30 copies during tissue culture (Hirochika *et al.*, 1996). More than 47 000 *Tos17* insertion lines have been produced for generating insertional mutants in rice (Hirochika, 2001; Miyao *et al.*, 2003).

A LINE-type retrotransposon, *Karma*, has been identified from rice as an insertion in a mutant allele of the *FRIZZY PANICLE2* locus (Komatsu *et al.*, 2003a). Its transcription *Karma* is activated in cultured cells upon DNA hypomethylation. Although the copy number of the element does not increase in either cultured cells or

the first generation of regenerated plants, it does rise in the next generations. Therefore, *Karma* may be used for producing insertional mutants in rice.

Activation tagging

Insertional mutagenesis usually generates recessive loss-of-function mutations, making them unsuitable for functional analysis of redundant genes. About two-thirds of the *Arabidopsis* genome is duplicated in the form of large chromosomal segments (The Arabidopsis Genome Initiative, 2000). In addition, about 4000 genes are tandemly repeated as two or more copies. Many rice genes are members of a multigene family (Goff *et al.*, 2002; Sasaki *et al.*, 2002; Yu *et al.*, 2002). Therefore, unless multiple mutants are generated, T-DNA tagging or transposon tagging proves inefficient in studying these genes. It is estimated that fewer than 10% of the genes tagged in *Arabidopsis* and rice are likely to generate a visible phenotypic change (Feldmann, 1991; Bouche and Bouchez, 2001; Jeon and An, 2001). Thus, complementing technologies are needed for studying genes whose function cannot be resolved by insertional mutagenesis.

Since activation tagging in plants was first proposed for studying gene function (Walden *et al.*, 1994), it has been effectively applied in *Arabidopsis* using T-DNA inserts (Kardailsky *et al.*, 1999; Ito and Meyerowitz, 2000; Lee *et al.*, 2000; Weigel *et al.*, 2000) and transposon systems (Wilson *et al.*, 1996; Marsch-Martinez *et al.*, 2002). These inserts carry strong activator elements that enhance the expression of genes adjacent to the insertion site. For example, the cauliflower mosaic virus (CaMV) 35S enhancer or promoter sequences (Odell *et al.*, 1985) have been used as system activators. Their primary function is to over-express tagged genes to reveal dominant gain-of-function phenotypes (Figure 2). In a number of cases, the enhancer acts by quantitatively increasing the original expression patterns, rather than by promoting ectopic or constitutive overexpression (Neff *et al.*, 1999; van der Graaff *et al.*, 2000). From the T-DNA activation-tagging pools of *Arabidopsis*, Weigel *et al.*, (2000) have characterized over 30 dominant mutants with various phenotypes. Analysis of a subset has shown that the tagging vector causes over-expression of the gene immediately adjacent to the inserted enhancer.

Development of activation tagging systems in rice requires a strong enhancer element that functions in this model monocot species. Jeong *et al.* (2002) have demonstrated that the CaMV 35S enhancer element efficiently increases expression of nearby genes from either their 5' or 3' ends. Based on this, they have devised a binary vector that carries the tetramerized 35S enhancers next to the left T-DNA border. This vector has now been used to generate activation-tagging pools for more than 50 000 individual transformants (Jeong *et al.*, 2002). Moreover, examination of randomly selected tags in the intergenic regions has shown that approximately 40% of the enhancer insertions increase nearby gene expression, proving that activation tagging efficiency at the transcriptional level is quite high. However, enhanced gene expression does not always result in phenotypic alteration because most activation taggings cause a quantitative increase in endogenous expression without changing tissue-specificity (Neff *et al.*, 1999; Weigel *et al.*, 2000; Jeong *et al.*, 2002). This differs from ectopic expression by strong promoters such as *actin* or *ubiquitin* promoters, which cause constitutive over-expression in all tissues (McElroy *et al.*, 1990; Cornejo *et al.*, 1993; LeClere and Bartel, 2001). Nonetheless, the dominant mutation frequency of an activation tagging population is much higher than that of a tag population generated by a simple insertion vector, thereby indicating that the activation tagging strategy works in rice. In fact, characterization of these dominant mutants has resulted in the identification of a number of genes tagged by the enhancer (Jeong *et al.*, unpublished data).

Entrapment tagging

Gene function can also be studied by elucidating its expression pattern. Analysis of transcript levels by RNA-gel blot or RT-PCR experiments provides only gross-level information on where the gene is expressed. Because RNA or antibody *in-situ* experiments are tedious and time consuming, they are unsuitable for large-scale analyses of many genes. These difficulties can be overcome by entrapment tagging. In this system, T-DNA or transposons are engineered to carry a promoterless reporter gene, such as β -glucuronidase (*GUS*) or *green fluorescent protein* (*GFP*), next to the insert

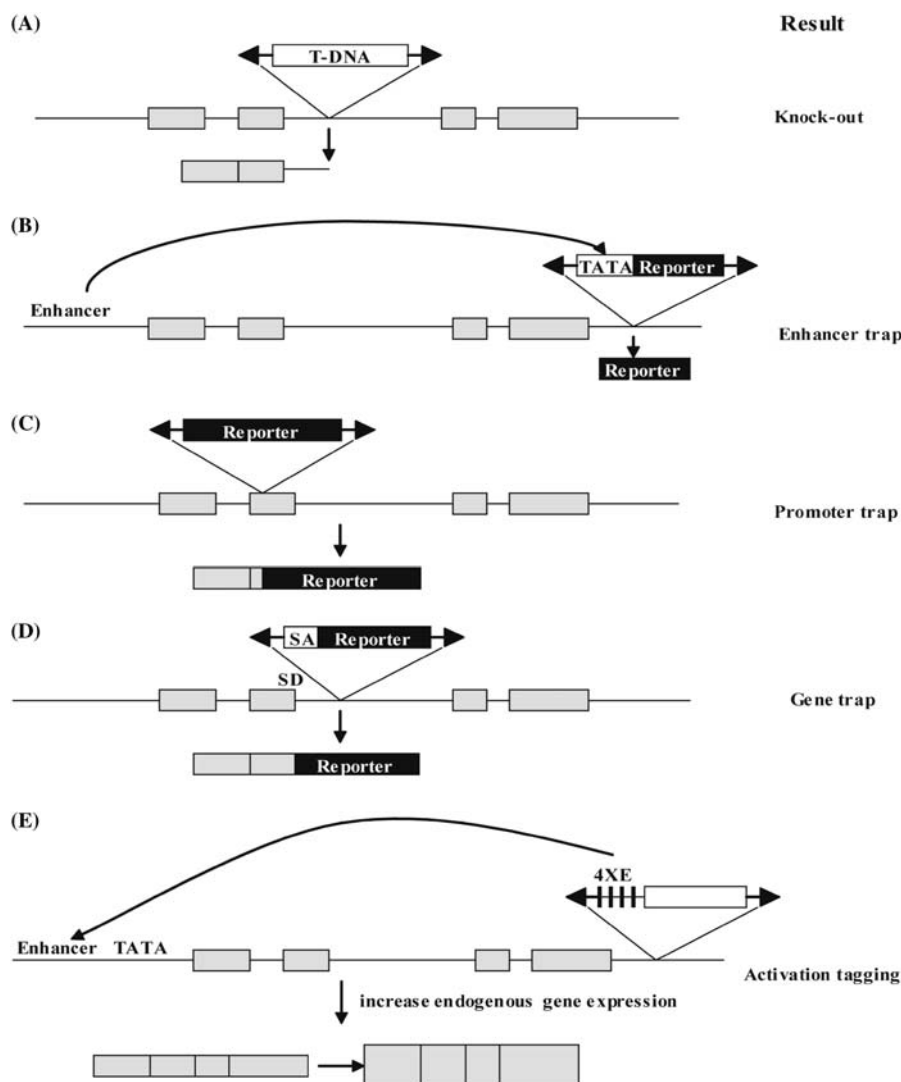


Figure 2. Schematic diagram of taggings. (A) Knock out. Simple T-DNA insertion in either exons (boxes) and introns (lines) can disrupt expression of the tagged gene. (B) Enhancer trap. The minimal promoter (TATA) of a reporter gene is activated by an enhancer element, resulting in expression of the reporter gene. (C) Promoter trap. The promoterless reporter gene can be expressed when insertion occurs in an exon so as to result in a transcriptional fusion. (D) Gene trap. The promoterless reporter gene contains splice acceptor (SA) sequences, which allow a transcriptional fusion between the tagged gene and the reporter gene by splicing from the splice donor (SD) site to the SA sequence. Therefore, in this system, the reporter can be activated even when it is inserted into an intron. (E) Activation tagging. Enhancer elements (4XE) of CaMV 35S promoter increase the expression of the gene near the insertion position.

border (Jefferson *et al.*, 1987; Pang *et al.*, 1996; Ryu *et al.*, 2004). Inserting the reporter into a gene not only destroys normal function but also activates expression of the reporter gene (Figure 2). Thus, expression of the tagged genes can be easily monitored by simple enzyme assays.

The *GUS* gene has frequently been used for entrapment tagging in plants because of the

accurate detection of its gene product and tolerance for the N-terminal translational fusions in its enzyme activity. However, one problem is the destructive nature of its assay staining and destaining procedures (Jefferson *et al.*, 1987). Therefore, non-invasive and non-destructive reporter genes, such as *GFP* or *luciferase*, are useful for entrapment in plants. Tagging popula-

tions with the *GFP* gene have been established for *Arabidopsis* and rice (Haseloff *et al.* 1997; Ryu *et al.*, 2004). With its rapid turnover and half-life of 3 h, luciferase is well suited as a real-time reporter for *in-planta* gene expression studies (Thompson *et al.*, 1991; Millar *et al.*, 1992). Recently, the enhancer trap system in tomato and *Arabidopsis* has demonstrated the usefulness of the reporter in random gene tagging experiments (Meissner *et al.*, 2000; Alvarado *et al.*, 2004).

Entrapment systems can be divided into three groups: enhancer trap, promoter trap, and gene trap (Springer, 2000) (Figure 2). In the enhancer trap system, the reporter gene is fused to a minimal promoter, which is unable to drive reporter gene expression but can be activated by an enhancer element of the neighboring genes (Klimyuk *et al.*, 1995; Campisi *et al.*, 1999).

Although enhancer trapping yields a higher frequency of *in-situ* reporter gene activation, promoter trapping and gene trapping have an additional feature that yields loss-of-function gene mutations. In the promoter trap system, reporter-gene expression can occur when the reporter is inserted into an exon and forms a translational fusion with the endogenous gene (Topping and Lindsey, 1997). In contrast, gene trap constructs have an intron with multiple splice donor and acceptor sequences in front of a reporter gene (Sundaresan *et al.*, 1995). This allows splicing from the donor sites in the disrupted gene to the acceptor sites in the reporter gene, resulting in fusion of the upstream exon sequences to the reporter. Therefore, the reporter gene can be expressed regardless of insert position (exon, intron, or UTR).

This approach has been successfully applied to genes that are difficult to identify by traditional methods. These include regulatory sequences that drive reporter gene expression in nematode feeding structures (Barthels *et al.*, 1997), molecular markers for embryogenesis (Topping and Lindsey, 1997), regulatory sequences that mediate guard cell-specific expression (Plesch *et al.*, 2000), enhancer trapped lines that show circadian-regulated expression (Michael and McClung, 2003), or senescence-associated expression (He *et al.*, 2001).

Activation of the reporter gene in a promoter trap vector can be as high as 30% (Sundaresan *et al.*, 1995). In rice, at least 5% of the T-DNAs and 10% of the transposed *Ds* elements become

activated in various tissues, e.g., roots, leaves, flowers, and seeds (Chin *et al.*, 1999; Jeon *et al.*, 2000). If one includes reporter gene activation by certain environmental conditions or by chemicals, such as growth substances, total tagging efficiency is higher. Some tags display tissue- or organ-specific reporter expression, while others exhibit ubiquitous expression patterns.

An enhancer trap system has been developed for rice (Greco *et al.*, 2003; Wu *et al.*, 2003). Specifically, a modified enhancer system uses the yeast transcription factor GAL4. There, the activator domain of GAL4 is replaced by the activation domain of Herpes simplex virus protein 16 (Triezenberg *et al.*, 1988). This GAL4/VP16 system has been applied to the construction of enhancer trap lines in rice (Wu *et al.*, 2003).

Forward screening

Insertional mutant lines are useful resources for studying gene function. However, most knockout mutations are recessive. Therefore, their phenotypes are not visible in the primary transgenic lines and can be detected only in the progeny generations. Because rice plants are relatively tall and wide, compared with *Arabidopsis*, forward screening of a mutant is more laborious and requires extra space. In their analyses of T-DNA tagging lines, Jeon *et al.* (2000) have observed that the most common mutants are dwarfing (7.0%) and leaf-pigment alterations (9.5%). Developmental mutants have also been observed in their screening. Likewise, Wu *et al.* (2003) have reported morphologically conspicuous mutations in about 7.5% of their 2679 T-DNA tagging lines.

One must determine whether these phenotypes are due to the insertion. Co-segregation of phenotypes with the insert DNA is a good indication that the phenotype is caused by the insert. Ultimately, the mutant phenotype must be confirmed by another allele or complemented by the wild-type gene. Tissue culturing is known to generate mutations not associated with insertions (Phillips *et al.*, 1994; Kaeppler *et al.*, 2000). Retrotransposons, such as *Tos17* or *Karma*, are activated during tissue culture (Hirochika *et al.*, 1996; Komatsu *et al.*, 2003b). Miniature inverted-repeat transposable elements (MITEs) e.g., *mPing* or *Pong*, can also act as mutagens (Jiang *et al.*, 2003;

Kikuchi *et al.*, 2003; Nakazaki *et al.*, 2003). In addition, small insertions, deletions, and base substitutions may be induced in cultured cells (Miyao *et al.*, 2003). Therefore, a large portion of mutants are generated in these populations through a mechanism not associated with the introduced inserts.

Nonetheless, insertional mutant populations are good resources for studying gene function with forward genetics. By analyzing chlorina mutants in rice, Jung *et al.* (2003) have isolated T-DNA and *Tos17* insertions into the *magnesium chelatase* gene. In addition, the study of pale green mutants has revealed T-DNA insertions in the *chlorophyll a oxygenase* gene (Lee *et al.*, unpublished). Agrawal *et al.* (2001) have screened for viviparous mutants and identified *Tos17* insertions in the rice *zeaxanthin epoxidase* gene (*OsABA1*) and in a novel *OsTATC* gene. Finally, the screening of brittle culm mutants has led to the identification of *Tos17* insertions in the *cellulose synthase* gene (Tanaka *et al.*, 2003).

DNA pool screening

Insertional mutant collections are more beneficial when used with reverse genetics approaches. With the near completion of rice genome sequencing, researchers are now identifying a large number of genes based on their sequence homology. Generally, groupings of related genes are formed, in which the functions of only one or a few have been elucidated. Reverse genetics approaches are excellent methods for studying a group of genes relatively quickly. In these approaches, mutants in a target gene are first collected, then their phenotypes are investigated. Double or multiple mutants may be generated as necessary for further assessment.

Insertional mutants in a given gene can be isolated via PCR-based screening (Krysan *et al.*, 1999; Sato *et al.*, 1999). Using a gene-specific primer and a primer located near the end of the insert, a DNA fragment flanking the insert element can be amplified and its sequence then determined. This strategy has already been successfully applied with *Arabidopsis*, petunia, and maize (Koes *et al.*, 1995; Mena *et al.*, 1996; Krysan *et al.*, 1999; Parinov and Sundaresan, 2000). For example, 17 insertions in 63 genes involved in signal transduction and ion transport,

47 insertions in 36 members of the *R2R3 MYB* gene family, and 22 mutations in 70 *P450* genes have been isolated from *Arabidopsis* (Winkler *et al.*, 1998; Krysan *et al.*, 1999; Meissner *et al.*, 1999). Screening of individual lines requires great effort, so DNA pools of a large number of lines are commonly used for this reverse genetics approach. In *Arabidopsis*, pools of 1000–5000 lines have been efficiently utilized for identifying insertional mutations (Winkler *et al.*, 1998; Krysan *et al.*, 1999; Meissner *et al.*, 1999; Rios *et al.*, 2002). Because the rice genome is about four times larger than that of *Arabidopsis*, a pool of 1000 lines should be sufficient for PCR-based screening of its knockout mutations (Lee *et al.*, 2003).

Rice has a higher amount of GC than does *Arabidopsis* (Carels and Bernardi, 2000; Yu *et al.*, 2002). In addition, rice shows a much higher GC value at the 5' end, which gradually decreases at the 3' end to a value slightly higher than that in *Arabidopsis* (Yu *et al.*, 2002). Because these high-GC regions are difficult to amplify under normal PCR conditions, including betaine in the buffer and adjusting the annealing temperature can improve PCR efficiency in those regions (Hengen, 1997; Henke *et al.*, 1997; Lee *et al.*, 2003).

Enoki *et al.* (1999) and Hirochika (2001) have reported PCR-based reverse genetics screenings of the *Ac* and *Tos17* insertion lines in rice. Of 14 randomly selected genes, two knockouts have been identified from *Ac* pools made from 6000 individuals (Enoki *et al.*, 1999). One knockout mutant in *OSH15* has also been isolated from 47 DNA pools that average 550 individuals each (Sato *et al.*, 1999). Three independent phytochrome-A mutants have been found from *Tos17* pools (Takano *et al.*, 2001). Finally, attempts to obtain knockout mutants in 12 MADS box genes from 21 049 T-DNA tagged individuals has resulted in the identification of five insertions in four genes (Lee *et al.*, 2003). A DNA pool screening service for T-DNA insertional lines in rice is now available to the scientific community (Lee *et al.*, 2003).

Tag end sequence (TES) database

Tag end sequences (TES) that flank insert elements have been obtained by TAIL PCR, iPCR, or

adaptor-ligation PCR (Ochman *et al.*, 1988; Triglia *et al.*, 1988; Rosenthal and Jones, 1990; Liu and Whittier, 1995; Siebert *et al.*, 1995). When numerous flanking sequences are generated, they are catalogued into databases (Parinov *et al.*, 1999; Tissier *et al.*, 1999; Parinov and Sundaresan, 2000). Although large-scale application of this alternative strategy requires considerable effort, once established, these databases can be easily shared with other scientists, facilitating the distribution of mutant materials and analysis of gene function. Because sequencing of their entire genomes is nearly complete, flanking-sequence databases for rice and *Arabidopsis* will become powerful tools for systematically analyzing the functions of a large number of genes in those species (Parinov and Sundaresan, 2000; Walbot, 1992, 2000; Kumar and Hirochika, 2001; Pan *et al.*, 2003). Databases of *Ds* transposon insertion site sequences and T-DNA insertion sites have already been established for *Arabidopsis* (Parinov *et al.*, 1999; Tissier *et al.*, 1999; Ortega *et al.*, 2002; Sessions *et al.*, 2002). In maize, the DNAs adjacent to transposed *Ac* elements have also been isolated and sequenced (Cowperthwaite *et al.*, 2002).

In rice, a TES database has been derived from the *Tos17* insertional mutant lines (Hirochika, 2001; Yamazaki *et al.*, 2001; Miyao *et al.*, 2003). Hirochika's group has analyzed more than 20 000 TES from 4316 mutant lines (tos.nais.affrc.go.jp). TES databases for T-DNA insertional lines have also been created. An *et al.* (2003) has reported the establishment of a database of 3793 TES (www.postech.ac.kr/life/pfg/risd), while Chen *et al.* (2003) has described one with 1009 TES (www.genomics.zju.edu.cn/ricetdna). TES analyses reveal the distribution of insert elements in plant chromosomes. For example, most transposition events occur in a tightly linked site, as reported with *Arabidopsis* (Sundaresan, 1996; Machida *et al.*, 1997). In that species, the physical map positions of 356 *Ds* insertions show significant preference for transposition to areas adjacent to the nucleolus organizer regions (Parinov *et al.*, 1999). Therefore, a short-range and highly preferential transposition system can be effectively utilized for targeted mutagenesis of closely linked genes.

For any given gene, *Ds* transposes equally into exons and introns, indicating no preference toward the 5' end for genes in *Arabidopsis*, maize, or rice (Parinov *et al.*, 1999; Cowperthwaite *et al.*, 2002;

Kolesnik *et al.*, 2004). Analysis of 2057 *Ds* flanking sequences in rice has demonstrated that *Ds* insertions are distributed randomly throughout that genome (Kolesnik *et al.*, 2004). However, a bias exists toward Chromosomes 4 and 7, which have twice as many insertions as expected. One *Ds* hot spot can be found within a 40-kbp region on Chromosome 7. Moreover, analysis of 200 *Ds* tags on Chromosome 1 has revealed that 72% of the insertions are located in a genic region. Miyao *et al.* (2003) have investigated more than 20 000 unique insertion points in the *Tos17* tagged sites and report that insertion events are three times more frequent in genic regions than in intergenic regions. Consistent with this, *Tos17* prefers gene-dense regions over centromeric heterochromatin regions. Although insertion target sites are distributed throughout the rice chromosomes, they tend to cluster. This phenomenon may reduce the efficiency of genome-wide gene disruption by *Tos17*.

Analysis of TES has demonstrated that the frequency of T-DNA insertion into genic regions is greater than the expected probability given random insertion (An *et al.*, 2003; Chen *et al.*, 2003). Frequencies are also higher at the beginning and end of the coding regions, as well as upstream near the start ATG codon. The overall GC content at the insertion sites is close to that measured for the entire rice genome. Functional classification of 1846 tagged genes has shown a distribution similar to that observed for all genes in the rice chromosomes. This indicates that T-DNA insertion is not biased toward a particular functional group (An *et al.*, 2003). At the chromosomal level, insertions are not evenly distributed; frequencies are higher at the ends and lower near the centromere. At certain sites, the frequency is greater than in the surrounding regions, proving that T-DNA insertion is not a random event.

Prospects

The number of insertional mutants accumulated thus far by the scientific community is large enough to saturate almost all the genes in japonica rice (Hirochika *et al.*, 2004). Because nearly all mutants have been generated in japonica varieties, it will be valuable to develop methods for tagging indica varieties. The improvement of transforma-

tion efficiency is a critical factor when generating a large number of transgenic plants. The *Ds* system has an advantage over T-DNA tagging because a smaller initial number of plants are needed. A method that utilizes an endogenous transposon means that a large number of inserts can be generated without going through the transformation procedure. *Tos17* or the MITE element *mPING*, apparently active in cultured indica cells, presents one possibility for generating insertional mutant populations in that way.

The generation of TES databases is also urgently needed for efficient utilization of this insertional mutant resource. DNA pools are useful for identifying the alleles of a given gene. Such reverse genetics materials are valuable to functional genomic analyses of a large number of genes in rice.

Acknowledgements

We thank Priscilla Licht for critical editing of the manuscript. The gene tagging project for functional genomics of rice in the An laboratory is supported, in part, by grants from the Crop Functional Genomic Center, the 21st Century Frontier Program (CG-1111), and from the Biogreen 21 Program, Rural Development Administration.

References

- Agrawal, G.K., Yamazaki, M., Kobayashi, M., Hirochika, R., Miyao, A. and Hirochika, H. 2001. Screening of the rice viviparous mutants generated by endogenous retrotransposon *Tos17* insertion. Tagging of a zeaxanthin epoxidase gene and a novel *OsTATC* gene. *Plant Physiol.* 125: 1248–1257.
- Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C. and Ecker, J.R. 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301: 653–657.
- Alvarado, M.C., Zsigmond, L.M., Kovacs, I., Cseplo, A., Koncz, C. and Szabados, L.M. 2004. Gene trapping with firefly luciferase in *Arabidopsis*. Tagging of stress-responsive genes. *Plant Physiol.* 134: 18–27.
- An, S., Park, S., Jeong, D.-H., Lee, D.Y., Kang, H.G., Yu, J.H., Hur, J., Kim, S.R., Kim, Y.H., Lee, M., Han, S., Kim, S.J., Yang, J., Kim, E., Wi, S.J., Chung, H.S., Hong, J.P., Choe, V., Lee, H.K., Choi, J.H., Nam, J., Kim, S.R., Park, P.B., Park, K.Y., Kim, W.T., Choe, S., Lee, C.B. and An, G. 2003. Generation and analysis of end sequence database for T-DNA tagging lines in rice. *Plant Physiol.* 133: 2040–2047.
- Azpiroz-Leehan, R. and Feldmann, K.A. 1997. T-DNA insertion mutagenesis in *Arabidopsis*: going back and forth. *Trends Genet.* 13: 152–156.
- Barakat, A., Gallois, P., Raynal, M., Mestre-Ortega, D., Sallaud, C., Guiderdoni, E., Delseny, M. and Bernardi, G. 2000. The distribution of T-DNA in the genomes of transgenic *Arabidopsis* and rice. *FEBS Lett.* 471: 161–164.
- Barthels, N., van der Lee, F.M., Klap, J., Goddijn, O.J., Karimi, M., Puzio, P., Grundler, F.M., Ohl, S.A., Lindsey, K., Robertson, L., Robertson, W.M., van Montagu, M., Gheysen, G. and Sijmons, P.C. 1997. Regulatory sequences of *Arabidopsis* drive reporter gene expression in nematode feeding structures. *Plant Cell* 9: 2119–2134.
- Bouche, N. and Bouchez, D. 2001. *Arabidopsis* gene knockout: phenotypes wanted. *Curr. Opin. Plant Biol.* 4: 111–117.
- Campisi, L., Yang, Y.Z., Yi, Y., Heilig, E., Herman, B., Cassista, A.J., Allen, D.W., Xiang, H.J. and Jack, T. 1999. Generation of enhancer trap lines in *Arabidopsis* and characterization of expression patterns in the inflorescence. *Plant J.* 17: 699–707.
- Carels, N. and Bernardi, G. 2000. Two classes of genes in plants. *Genetics* 154: 1819–1825.
- Chen, S., Jin, W., Wang, M., Zhang, F., Zhou, J., Jia, Q., Wu, Y., Liu, F. and Wu, P. 2003. Distribution and characterization of over 1000 T-DNA tags in rice genome. *Plant J.* 36: 105–113.
- Chin, H.G., Choe, M.S., Lee, S.H., Park, S.H., Koo, J.C., Kim, N.Y., Lee, J.J., Oh, B.G., Yi, G.H., Kim, S.C., Choi, H.C., Cho, M.J. and Han, C.D. 1999. Molecular analysis of rice plants harboring an *Ac/Ds* transposable element-mediated gene trapping system. *Plant J.* 19: 615–623.
- Chuang, C.F. and Meyerowitz, E.M. 2000. Specific and heritable genetic interference by doublestranded RNA in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 97: 4985–4990.
- Cornejo, M.J., Luth, D., Blankenship, K.M., Anderson, O.D. and Blechl, A.E. 1993. Activity of a maize ubiquitin promoter in transgenic rice. *Plant. Mol. Biol.* 23: 567–581.
- Cowperthwaite, M., Park, W., Xu, Z., Yan, X., Maurais, S.C. and Dooner, H.K. 2002. Use of the transposon *Ac* as a gene-searching engine in the maize genome. *Plant Cell* 14: 713–726.
- Enoki, H., Izawa, T., Kawahara, M., Komatsu, M., Koh, S., Kyozuka, J. and Shimamoto K. 1999. *Ac* as a tool for the functional genomics of rice. *Plant J.* 19: 605–613.
- Fedoroff, N.V., Wessler, S. and Shure, M. 1983. Isolation of the transposable maize controlling elements *Ac* and *Ds*. *Cell* 35: 235–242.
- Feldmann, K.A. 1991. T-DNA insertion mutagenesis in *Arabidopsis*: mutational spectrum. *Plant J.* 1: 71–82.
- Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchison, D., Martin, C., Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W.L., Chen, L., Cooper, B., Park, S., Wood, T.C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, S., Mitchell, J.,

- Eldredge, G., Scholl, T., Miller, R.M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus, J., Macalma, T., Oliphant, A. and Briggs, S. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296: 92–100.
- Greco, R., Ouwerkerk, P.B., de Kam, R.J., Sallaud, C., Favalli, C., Colombo, L., Guiderdoni, E., Meijer, A.H., Hoge Dagger, J.H. and Pereira, A. 2003. Transpositional behaviour of an *Ac/Ds* system for reverse genetics in rice. *Theor. Appl. Genet.* 108: 10–24.
- Hanin, M. and Paszkowski, J. 2003. Plant genome modification by homologous recombination. *Curr. Opin. Plant Biol.* 6: 157–162.
- Haseloff, J., Siemering, K.R., Prasher, D.C. and Hodge, S. 1997. Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. *Proc. Natl. Acad. Sci. U.S.A.* 94: 2122–2127.
- He, Y., Tang, W., Swain, J.D., Green, A.L., Jack, T.P. and Gan, S. 2001. Networking senescence-regulating pathways by using *Arabidopsis* enhancer trap lines. *Plant Physiol.* 126: 707–716.
- Hengen, P.N. 1997. Optimizing multiplex and LA-PCR with betaine. *Trends Biochem. Sci.* 22: 225–226.
- Henke, W., Herdel, K., Jung, K., Schnorr, D. and Loening, S.A. 1997. Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucl. Acids Res.* 25: 3957–3958.
- Hiei, Y., Ohta, S., Komari, T. and Kumashiro, T. 1994. Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *Plant J.* 6: 271–282.
- Hirochika, H. 2001. Contribution of the *Tos17* retrotransposon to rice functional genomics. *Curr. Opin. Plant Biol.* 4: 118–122.
- Hirochika, H., Guiderdoni, E., An, G., Hsing, Y., Eun, M.Y., Han, C., Upadhyaya, N., Ramachandran, S., Zhang, Q., Pereira, A., Sundaresan, V., and Leung, H. 2004. Rice mutant resources for gene discovery. *Plant Mol. Biol.* 54: 325–334.
- Hirochika, H., Sugimoto, K., Otsuki, Y., Tsugawa, H. and Kanda, M. 1996. Retrotransposons of rice involved in mutations induced by tissue culture. *Proc. Natl. Acad. Sci. U.S.A.* 93: 7783–7788.
- Ito, T. and Meyerowitz, E.M. 2000. Overexpression of a gene encoding a cytochrome P450, *CYP78A9*, induces large and seedless fruit in *Arabidopsis*. *Plant Cell* 12: 1541–1550.
- Izawa, T., Miyazaki, C., Yamamoto, M., Terada, R., Iida, S. and Shimamoto, K. 1991. Introduction and transposition of the maize transposable element *Ac* in rice (*Oryza sativa* L.). *Mol. Gen. Genet.* 227: 391–396.
- Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. 1987. GUS fusions: *beta-glucuronidase* as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* 6: 3901–3907.
- Jeon, J.S. and An, G. 2001. Gene tagging in rice: a high throughput system for functional genomics. *Plant Sci.* 161: 211–219.
- Jeon, J.S., Lee, S., Jung, K.H., Jun, S.H., Jeong, D.-H., Lee, J., Kim, C., Jang, S., Yang, K., Nam, J., An, K., Han, M.J., Sung, R.J., Choi, H.S., Yu, J.H., Choi, J.H., Cho, S.Y., Cha, S.S., Kim, S.I. and An, G. 2000. T-DNA insertional mutagenesis for functional genomics in rice. *Plant J.* 22: 561–570.
- Jeong, D.-H., An, S., Kang, H.-G., Moon, S., Han, J.-J., Park, S., Lee, H.S., An, K. and An, G. 2002. T-DNA insertional mutagenesis for activation tagging in rice. *Plant Physiol.* 130: 1636–1644.
- Jiang, N., Bao, Z., Zhang, X., Hirochika, H., Eddy, S.R., McCouch, S.R. and Wessler, S.R. 2003. An active DNA transposon family in rice. *Nature* 421: 163–167.
- Jordan, I.K., Rogozin, I.B., Glazko, G.V. and Koonin, E.V. 2003. Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet.* 19: 68–72.
- Jung, K.H., Hur, J., Ryu, C.H., Choi, Y., Chung, Y.Y., Miyao, A., Hirochika, H. and An, G. 2003. Characterization of a rice chlorophyll-deficient mutant using the T-DNA gene-trap system. *Plant Cell Physiol.* 33: 463–472.
- Kaeppler, S.M., Kaeppler, H.F. and Rhee, Y. 2000. Epigenetic aspects of somaclonal variation in plants. *Plant Mol. Biol.* 43: 179–188.
- Kardailsky, I., Shukla, V.K., Ahn, J.H., Dagenais, N., Christensen, S.K., Nguyen, J.T., Chory, J., Harrison, M.J. and Weigel, D. 1999. Activation tagging of the floral inducer *FT*. *Science* 286: 1962–1965.
- Kashkush, K., Feldman, M. and Levy, A.A. 2003. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat. Genet.* 33: 102–106.
- Kikuchi, K., Terauchi, K., Wada, M. and Hirano, H.Y. 2003. The plant MITE mPing is mobilized in anther culture. *Nature* 421: 167–170.
- Kim, C.M., Je, B.I., Piao, H.L., Park, S.J., Kim, M.J., Park, S.H., Park, J.Y., Park, S.H., Lee, E.K., Chon, N.S., Won, Y.J., Lee, G.H., Nam, M.H., Yun, D.W., Lee, M.C., Cha, Y.S., Lee, K.H., Eun, M.Y. and Han, C.D. 2002. Reprogramming of the activity of the activator/dissociation transposon family during regeneration in rice. *Mol. Cells* 14: 231–237.
- Klimyuk, V.I., Nussaume, L., Harrison, K. and Jones, J.D.G. 1995. Novel *gus* expression patterns following transposition of an enhancer trap *Ds* element in *Arabidopsis*. *Mol. Gen. Genet.* 249: 357–365.
- Koes, R., Souer, E., van Houwelingen, A., Mur, L., Spelt, C., Quattrocchio, F., Wing, J., Oppedijk, B., Ahmed, S., Maes, T., Gerats, T., Hoogeveen, P., Meesters, M., Kloos, D. and Mol, J.N.M. 1995. Targeted gene inactivation in petunia by PCR-based selection of transposon insertion mutants. *Proc. Natl. Acad. Sci. U.S.A.* 92: 8149–8153.
- Kolesnik, T., Szeverenyi, I., Bachmann, D., Kumar, C.S., Jiang, S., Ramamoorthy, R., Cai, M., Ma, Z.G., Sundaresan, V. and Ramachandran, S. 2004. Establishing an efficient *Ac/Ds* tagging system in rice: large-scale analysis of *Ds* flanking sequences. *Plant J.* 37: 301–314.
- Komatsu, M., Chujo, A., Nagato, Y., Shimamoto, K. and Kyozuka, J. 2003a. FRIZZY PANICLE is required to prevent the formation of axillary meristems and to establish floral meristem identity in rice spikelets. *Development* 130: 3841–3850.
- Komatsu, M., Shimamoto, K. and Kyozuka, J. 2003b. Two-step regulation and continuous retrotransposition of the rice LINE-type retrotransposon *Karma*. *Plant Cell* 15: 1934–1944.
- Krysan, P.J., Young, J.C. and Sussman, M.R. 1999. T-DNA as an insertional mutagen in *Arabidopsis*. *Plant Cell* 11: 2283–2290.

- Kumar, S. and Fladung, M. 2000. Transgene repeats in aspen: molecular characterization suggests simultaneous integration of independent T-DNA into receptive hotspots in the host genome. *Mol. Gen. Genet.* 264: 20–28.
- Kumar, A. and Hirochika, H. 2001. Applications of retrotransposons as genetic tools in plant biology. *Trends Plant Sci.* 6:127–134.
- LeClere, S. and Bartel, B. 2001. A library of *Arabidopsis* 35S-cDNA lines for identifying novel mutants. *Plant Mol. Biol.* 46: 695–703.
- Lee, H., Suh, S.S., Park, E., Cho, E., Ahn, J.H., Kim, S.G., Lee, J.S., Kwon, Y.M. and Lee, I. 2000. The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in *Arabidopsis*. *Genes Dev.* 14: 2366–2376.
- Lee, S., Jeon, J.S., Jung, K.H. and An, G. 1999. Binary vectors for efficient transformation of rice. *J. Plant Biol.* 42: 310–316.
- Lee, S., Kim, J., Son, J.S., Nam, J., Jeong, D.-H., Lee, K., Jang, S., Yoo, J., Lee, J., Lee, D.Y., Kang, H.G. and An, G. 2003. Systematic reverse genetic screening of T-DNA tagged genes in rice for functional genomic analyses: MADS-box genes as a test case. *Plant Cell Physiol.* 44: 1403–1411.
- Liu, Y.G. and Whittier, R. 1995. Thermal asymmetric inter-laced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* 25: 674–681.
- Machida, C., Onouchi, H., Koizumi, J., Hamada, S., Semiarti, E., Torikai, S. and Machida, Y. 1997. Characterization of the transposition pattern of the *Ac* element in *Arabidopsis thaliana* using endonuclease I-SceI. *Proc. Natl. Acad. Sci. U.S.A.* 94: 8675–8680.
- Marsch-Martinez, N., Greco, R., van Arkel, G., Herrera-Estrella, L. and Pereira, A. 2002. Activation tagging using the *En-1* maize transposon system in *Arabidopsis*. *Plant Physiol.* 129: 1544–1556.
- McCarthy, E.M., Liu, J., Lizhi, G. and McDonald, J.F. 2002. Long terminal repeat retrotransposons of *Oryza sativa*. *Genome Biol.* 3: 1–11.
- McElroy, D., Zhang, W., Cao, J. and Wu, R. 1990. Isolation of an efficient actin promoter for use in rice transformation. *Plant Cell* 2: 163–171.
- Meissner, R.C., Jin, H., Cominelli, E., Denekamp, M., Fuertes, A., Greco, R., Kranz, H.D., Penfield, S., Petroni, K., Urzainqui, A., Martin, C., Paz-Ares, J., Smeekens, S., Tonelli, C., Weisshaar, B., Baumann, E., Klimyuk, V., Marillonnet, S., Patel, K., Speulman, E., Tissier, A.F., Bouchez, D., Jones, J.J., Pereira, A., Wisman, E. and Bevan, M. 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.
- Meissner, R., Chague, V., Zhu, Q., Emmanuel, E., Elkind, Y. and Levy, A.A. 2000. Technical advance: a high throughput system for transposon tagging and promoter trapping in tomato. *Plant J.* 22: 265–274.
- Mena, M., Ambrose, B.A., Meeley, R.B., Briggs, S.P., Yanofsky, M.F. and Schmidt, R.J. 1996. Diversification of C-function activity in maize flower development. *Science* 274: 1537–1540.
- Michael, T.P. and McClung, C.R. 2003. Enhancer trapping reveals widespread circadian clock transcriptional control in *Arabidopsis*. *Plant Physiol.* 132: 629–639.
- Millar, A.J., Short, S.R., Chua, N.H. and Kay, S.A. 1992. A novel circadian phenotype based on firefly luciferase expression in transgenic plants. *Plant Cell* 4: 1075–1087.
- Miyao, A., Tanaka, K., Murata, K., Sawaki, H., Takeda, S., Abe, K., Shinozuka, Y., Onosato, K. and Hirochika, H. 2003. Target site specificity of the *Tos17* retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. *Plant Cell* 15: 1771–1780.
- Murai, N., Li, Z.J., Kawagoe, Y. and Hayashimoto, A. 1991. Transposition of the maize Activator element in transgenic rice plants. *Nucl. Acids Res.* 19: 617–622.
- Nakagawa, Y., Machida, C., Machida, Y. and Toriyama, K. 2000. Frequency and pattern of transposition of the maize transposable element *Ds* in transgenic rice plants. *Plant Cell Physiol.* 41: 733–742.
- Nakazaki, T., Okumoto, Y., Horibata, A., Yamahira, S., Teraishi, M., Nishida, H., Inoue, H. and Tanisaka, T. 2003. Mobilization of a transposon in the rice genome. *Nature* 421: 170–172.
- Neff, M.M., Nguyen, S.M., Malancharuvil, E.J., Fujioka, S., Noguchi, T., Seto, H., Tsubuki, M., Honda, T., Takatsuto, S., Yoshida, S. and Chory, J. 1999. *BAST1*: a gene regulating brassinosteroid levels and light responsiveness in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 96: 15316–15323.
- Nekrutenko, A. and Li, W. 2001. Transposable elements are found in a large number of human protein-coding genes. *Trends Genet.* 17: 619–621.
- Ochman, H., Gerber, A.S. and Hartl, D.L. 1988. Genetic applications of an inverse polymerase chain reaction. *Genetics* 120: 621–623.
- Odell, J.T., Nagy, F. and Chua, N. 1985. Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature* 313: 810–812.
- Ortega, D., Raynal, M., Laudie, M., Llauro, C., Cooke, R., Devic, M., Genestier, S., Picard, G., Abad, P., Contard, P., Sarrobert, C., Nussaume, L., Bechtold, N., Horlow, C., Pelletier, G. and Delseny, M. 2002. Flanking sequence tags in *Arabidopsis thaliana* T-DNA insertion lines: a pilot study. *C. R. Biol.* 325: 773–780.
- Pan, X., Liu, H., Clarke, J., Jones, J., Bevan, M. and Stein, L. 2003. ATIDB: *Arabidopsis thaliana* insertion database. *Nucl. Acids Res.* 31: 1245–1251.
- Pang, S.Z., DeBoer, D.L., Wan, Y., Ye, G., Layton, J.G., Neher, M.K., Armstrong, C.L., Fry, J.E., Hinchey, M.A. and Fromm, M.E. 1996. An improved green fluorescent protein gene as a vital marker in plants. *Plant Physiol.* 112: 893–900.
- Parinov, S., Sevugan, M., De, Y., 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.
- Parinov, S. and Sundaresan, V. 2000. Functional genomics in *Arabidopsis*: large-scale insertional mutagenesis complements the genome sequencing project. *Curr. Opin. Biotechnol.* 11: 157–161.
- Phillips, R.L., Kaeppeler, S.M. and Olhoft, P. 1994. Genetic instability of plant tissue cultures: breakdown of normal controls. *Proc. Natl. Acad. Sci. U.S.A.* 91: 5222–5226.
- Plesch, G., Kamann, E. and Mueller-Roeber, B. 2000. Cloning of regulatory sequences mediating guard-cell-specific gene expression. *Gene* 249: 83–89.
- Rios, G., Lossow, A., Hertel, B., Breuer, F., Schaefer, S., Broich, M., Kleinow, T., Jasik, J., Winter, J., Ferrando, A., Farras, R., Panicot, M., Henriques, R., Mariaux, J.-B., Oberschall, A., Molnar, G., Nerendzen, K., Shukla, V., Lafos, M., Koncz, Z., Redei, G.P., Schell, J. and Koncz, C.

2002. Rapid identification of *Arabidopsis* insertion mutants by non-radioactive detection of T-DNA tagged genes. *Plant J.* 32: 243–253.
- Rosenthal, A. and Jones, D.S. 1990. Genomic walking and sequencing by oligo-cassette mediated polymerase chain reaction. *Nucl. Acids Res.* 18: 3095–3096.
- Ryu, C.-H., You, J.-H., Kang, H.-G., Hur, J., Kim, Y.-H., Han, M.-J., An, K., Chung, B.-C., Lee, C.-H. and An, G., 2004. Generation of T-DNA tagging lines using a bi-directional gene trap vector and the establishment of an insertion-site database. *Plant Mol. Biol.* 54: 489–502.
- Sallaud, C., Meynard, D., van Bostel, J., Gay, C., Bes, M., Brizard, J.P., Larmande, P., Ortega, D., Raynal, M., Portefaix, M., Ouwerkerk, P.B., Rueb, S., Delseny, M. and Guiderdoni, E. 2003. Highly efficient production and characterization of T-DNA plants for rice (*Oryza sativa* L.) functional genomics. *Theor. Appl. Genet.* 106: 1396–1408.
- Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T., Katayose, Y., Wu, J., Niimura, Y., Cheng, Z., Nagamura, Y., Antonio, B.A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K., Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K., Hamada, M., Harada, C., Hijishita, S., Honda, M., Ichikawa, Y., Ido-numa, A., Iijima, M., Ikeda, M., Ikeno, M., Ito, S., Ito, T., Ito, Y., Ito, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S., Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T., Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M., Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M., Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T., Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Tsuji, K., Waki, K., Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K., Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J.H., Kim, H.-I., Eun, M.-Y., Yano, M., Jiang, J. and Gojobori, T. 2002. The genome sequence and structure of rice chromosome 1. *Nature* 420: 312–316.
- Sato, Y., Sentoku, N., Miura, Y., Hirochika, H., Kitano, H. and Matsuoka, M. 1999. Loss-of-function mutations in the rice homeobox gene *OSH15* affect the architecture of internodes resulting in dwarf plants. *EMBO J.* 18: 992–1002.
- Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B., Mitzel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S.A. 2002. A high-throughput *Arabidopsis* reverse genetics system. *Plant Cell* 14: 2985–2994.
- Sha, Y., Li, S., Pei, Z., Luo, L., Tian, Y. and He, C. 2004. Generation and flanking sequence analysis of a rice T-DNA tagged population. *Theor. Appl. Genet.* 108: 306–314.
- Shimamoto, K., Miyazaki, C., Hashimoto, H., Izawa, T., Itoh, K., Terada, R., Inagaki, Y. and Iida, S. 1993. Trans-activation and stable integration of the maize transposable element *Ds* cotransfected with the *Ac* transposase gene in transgenic rice plants. *Mol. Gen. Genet.* 239: 354–360.
- Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A. and Lukyanov, S.A. 1995. An improved PCR method for walking in uncloned genomic DNA. *Nucl. Acids Res.* 23: 1087–1088.
- Springer, P.S. 2000. Gene traps: tools for plant development and genomics. *Plant Cell* 12: 1007–1020.
- Sugimoto, K., Otsuki, Y., Saji, S. and Hirochika, H. 1994. Transposition of the maize *Ds* element from a viral vector to the rice genome. *Plant J.* 5: 863–871.
- Sundaresan, V. 1996. Horizontal spread of transposon mutagenesis: new uses of old elements. *Trends Plant Sci.* 1: 184–191.
- 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 transposable element. *Genes Dev.* 9: 1797–1810.
- Takano, M., Kanegae, H., Shinomura, T., Miyao, A., Hirochika, H. and Furuya, M. 2001. Isolation and characterization of rice phytochrome A mutants. *Plant Cell* 13: 521–534.
- Takeda, S., Sugimoto, K., Otsuki, H. and Hirochika, H. 1998. Transcriptional activation of the tobacco retrotransposon *Tto1* by wounding and methyl jasmonate. *Plant Mol. Biol.* 36: 365–376.
- Tanaka, K., Murata, K., Yamazaki, M., Onosato, K., Miyao, A. and Hirochika, H. 2003. Three distinct rice cellulose synthase catalytic subunit genes required for cellulose synthesis in the secondary wall. *Plant Physiol.* 133: 73–83.
- The Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815.
- Thompson, J.F., Hayes, L.S. and Lloyd, D.B. 1991. Modulation of firefly luciferase stability and impact on studies of gene regulation. *Gene* 103: 171–177.
- Tissier, A.F., Marillonnet, S., Klimyuk, V., Patel, K., Torres, M.A., Murphy, G. and Jones, J.D. 1999. Multiple independent defective suppressor-mutator transposon insertions in *Arabidopsis*: a tool for functional genomics. *Plant Cell* 11: 1841–1852.
- Topping, J.F. and Lindsey, K. 1997. Promoter trap markers differentiate structural and positional components of polar development in *Arabidopsis*. *Plant Cell* 9: 1713–1725.
- Triezenberg, S.J., Kingsbury, R.C. and McKnight, S.L. 1988. Functional dissection of VP16 the trans-activator of herpes simplex virus immediate early gene expression. *Gene Dev.* 2: 718–729.
- Triglia, T., Peterson, M.J. and Kemp, D.J. 1988. A procedure for *in vitro* amplification of DNA segments that lie outside the boundaries of known sequences. *Nucl. Acids Res.* 16: 8186.
- Upadhyaya, N.M., Zhou, X.-R., Zhu, Q.-H., Ramm, K., Wu, L., Eamens, A., Sivakumar, R., Kato, T., Yun, D.-W., Santhoshkumar, C., Narayanan, K.K., Peacock, J.W. and Dennis, E.S. 2002. An *iAc/Ds* gene and enhancer trapping system for insertional mutagenesis in rice. *Funct. Plant Biol.* 29: 547–559.
- van der Graaff, E., Dulk-Ras, A.D., Hooykaas, P.J. and Keller, B. 2000. Activation tagging of the *LEAFY PETIOLE* gene affects leaf petiole development in *Arabidopsis thaliana*. *Development* 127: 4971–4980.
- Walbot, V. 1992. Reactivation of *Mutator* transposable elements of maize by ultraviolet light. *Mol. Gen. Genet.* 234: 353–360.
- Walbot, V. 2000. Saturation mutagenesis using maize transposons. *Curr. Opin. Plant Biol.* 3: 103–107.
- Walden, R., Fritze, K., Hayashi, H., Miklashevichs, E., Harling, H. and Schell, J. 1994. Activation tagging: a means of isolating genes implicated as playing a role in plant growth and development. *Plant Mol. Biol.* 26: 1521–1528.
- Weigel, D., Ahn, J.H., Blazquez, M.A., Borevitz, J.O., Christensen, S.K., Fankhauser, C., Ferrandiz, C., Kardailsky, I., Malancharuvil, E.J., Neff, M.M., Nguyen, J.T., Sato, S., Wang, Z.Y., Xia, Y., Dixon, R.A., Harrison, M.J.,

- Lamb, C.J., Yanofsky, M.F. and Chory, J. 2000. Activation tagging in *Arabidopsis*. *Plant Physiol.* 122: 1003–1013.
- Wilson, K., Long, D., Swinburne, K. and Coupland, G. 1996. A *Dissociation* insertion causes a semidominant mutation that increases expression of *TINY*, an *Arabidopsis* gene related to *APETALA2*. *Plant Cell* 8: 659–671.
- Winkler, R.G., Frank, M.R., Galbraith, D.W., Feyereisen, R. and Feldmann, K.A. 1998. Systematic reverse genetics of transfer-DNA-tagged lines of *Arabidopsis*. Isolation of mutations in the cytochrome p450 gene superfamily. *Plant Physiol.* 118: 743–750.
- Wu, C., Li, X., Yuan, W., Chen, G., Kilian, A., Li, J., Xu, C., Li, X., Zhou, D.X., Wang, S. and Zhang, Q. 2003. Development of enhancer trap lines for functional analysis of the rice genome. *Plant J.* 35: 418–427.
- Yamazaki, M., Tsugawa, H., Miyao, A., Yano, M., Wu, J., Yamamoto, S., Matsumoto, T., Sasaki, T. and Hirochika, H. 2001. The rice retrotransposon *Tos17* prefers low-copy-number sequences as integration targets. *Mol. Genet. Genom.* 265: 336–344.
- Yin, Z. and Wang, G.-L. 2000. Evidence of multiple complex patterns of T-DNA integration into rice genome. *Theor. Appl. Genet.* 100: 461–470.
- Yu, J., Hu, S., Wang, J., Wong, G.K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., Sun, J., Tang, J., Chen, Y., Huang, X., Lin, W., Ye, C., Tong, W., Cong, L., Geng, J., Han, Y., Li, L., Li, W., Hu, G., Huang, X., Li, W., Li, J., Liu, Z., Li, L., Liu, J., Qi, Q., Liu, J., Li, L., Li, T., Wang, X., Lu, H., Wu, T., Zhu, M., Ni, P., Han, H., Dong, W., Ren, X., Feng, X., Cui, P., Li, X., Wang, H., Xu, X., Zhai, W., Xu, Z., Zhang, J., He, S., Zhang, J., Xu, J., Zhang, K., Zheng, X., Dong, J., Zeng, W., Tao, L., Ye, J., Tan, J., Ren, X., Chen, X., He, J., Liu, D., Tian, W., Tian, C., Xia, H., Bao, Q., Li, G., Gao, H., Cao, T., Wang, J., Zhao, W., Li, P., Chen, W., Wang, X., Zhang, Y., Hu, J., Wang, J., Liu, S., Yang, J., Zhang, G., Xiong, Y., Li, Z., Mao, L., Zhou, C., Zhu, Z., Chen, R., Hao, B., Zheng, W., Chen, S., Guo, W., Li, G., Liu, S., Tao, M., Wang, J., Zhu, L., Yuan, L. and Yang, H. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). *Science* 296: 79–92.