Abstract

Transposable elements are powerful mutagens. Along with genomic sequences, knock-out phenotypes and expression patterns are important information to elucidate the function of genes. In this review, I propose a strategy to develop transposant lines on a large scale by combining genetic cross and tissue culture of \( \text{Ac} \) and \( \text{Ds} \) lines. Based on the facts that \( \text{Ds} \) tends to be inactive in F2 or later generation and \( \text{Ds} \) becomes reactivated via tissue culture, a large scale of transposants can be produced by tissue culture of seeds carrying \( \text{Ac} \) and inactive \( \text{Ds} \). In this review, I describe limitations and considerations in operating transposon tagging systems in rice. Also, I discuss the efficiency of our gene trap system and technical procedures to clone \( \text{Ds} \) flanking DNA.

Rice genes and functional genomics

Two draft sequences of over 90 percent of the genomes of the \text{japonica} and \text{indica} rice varieties have been published (Goff et al., 2002; Yu et al., 2002) and the International Rice Genome Sequencing Project (IRGSP) will complete high quality sequencing of the rice genome in the near future (Leach et al., 2002). The gene number of rice is predicted to be 35 000 to 50 000, which is much more than that of \text{Arabidopsis}. Noticeably, many predicted rice genes are not found in the genomic sequence of \text{Arabidopsis}. This could be either due to complex chromosomal or genic duplications or to acquisition of monocot-specific genes. However, current bioinformatic programs suffer difficulty in performing accurate gene prediction from rice genomic sequence, mainly due to bias in base composition of coding regions. More time is needed to get more accurate genetic information from genomic sequence. To identify rice genes and understand the function of rice genes, it is essential to develop functional analysis tools. For functional genomics, insertional mutagenesis has been the most powerful strategy.

Transposable elements as mutagenizing agents in rice

Traditionally, in maize, where endogenous transposable elements have been well characterized, transposable element is a major tool to clone a gene that has only phenotypic information (Walbot, 2000). In other plants where transformation is a relatively easy procedure, enough to develop a large population, T-DNA and transposable element families (\( \text{Ac}/\text{Ds} \) or \( \text{Spm}/\text{dspm} \)) of maize have been the main insertional elements. Comparison of T-DNA and transposable
elements as insertional agent has been discussed elsewhere (Parinov and Sundaresan, 2000). Over 25 percent genome of rice carry structures related to transposable elements. However, one percent of genomic DNA was detected by cDNA via EST database. Many common elements of conserved structure, including MuDR, are detected in rice genome. However, there has been no report on an endogenous transposable element whose behavior during normal rice development is understood enough to be utilized for gene tagging system. In the case of retro-transposable elements, only under prolonged tissue culture stress, active mobility has been observed (Hirochika, 1997). In rice, Ac has been utilized as gene tagging vehicle (Izawa et al., 1997). However, there is no gene report on the success of another well-known element, Spm/En, in rice.

To identify rice genes and understand the function of rice genes, it is essential to develop functional analysis tools. For functional genomics, insertional mutagenesis has been the most powerful strategy.

Plants might be bred either carrying multi-copy of elements or containing single copy of transposable element. For development of multi-copy population, a transposable element is to be multiplied and maintained throughout many generations. MuDR of maize is a typical example that can be maintained in multi-copy in a genome. As a heterogenous element, Spm has been propagated and maintained in up to 20 copies in Arabidopsis (Wisman et al., 1998). In rice, retroposon, Tos elements have been proven to be multiplied in high copy. In maize, the single copy mutagenesis was done with Ac, which was shown to be very sensitive to the copy number and showed a negative dosage effect in maize. Since the transposition behavior is well studied in maize and Arabidopsis, Ac is a tool of choice for directed or targeted mutagenesis. However, it is still unclear whether Ac or Ds can be maintained in multi-copy in rice genome. I will discuss this topic in the last section.

**Gene trap Ds system in rice**

Gene trap is an advanced version of conventional transposon used for insertional mutagenesis (Skarnes, 1990). A gene trap transposon contains an intron with artificial splicing acceptor sites fused to a reporter gene in three different frames. A fusion protein of the reporter gene with the N-terminal portion of a host gene will be produced if the element is inserted into either an exon or an intron of the host gene in the same orientation of transcription. Such transposon trap systems have been successfully applied in Arabidopsis. Maize transposable elements Ac and Ds were engineered such that insertion of Ds carrying a reporter gene (GUS) into a cellular gene resulted in expression of the reporter gene (Sundaresan et al., 1995). This system is aimed to mobilize a single copy of Ds element by Ac, through genetic cross between Ac and Ds lines.

The following is the brief summary of our work on Ds trap system that has been established and evaluated during the last five years.

1) **Ac and Ds starter lines**

A two-element system has been adapted in generating transposant lines, in which immobile Ac supplies transposase to mobilize a non-autonomous Ds element containing a GUS reporter gene. Ac and Ds T-DNA vector were separately introduced into rice via Agrobacterium. CaMV 35S promoter was used to express Ac cDNA. Two gene trap Ds have been introduced into rice genomes. Two different introns with three alternative splicing acceptors were fused to GUS, and a BAR gene were inserted into gDs. The utilization efficiency of alternative splicing acceptors in gene trap Ds was examined using Agrobacterium-infected embryonic calli. Three cryptic splicing donor sites at the 3' end of Ds and two artificial splicing acceptor sites of GUS have been proven to be equally utilized in rice (Chin et al., 1999).

2) **Methods to develop Ds transposant lines**

The following three strategies have been employed to develop a large population of Ds transposants. These are 1) genetic crosses between Ac and Ds “starter” lines and selection of germinally transposed Ds in F2 generation, 2) mass plant regeneration, via tissue culture, from seeds carrying both Ac and Ds, 3) propagation and maintenance of multi-copy Ds lines.

a) **Genetic crosses between Ac and Ds “starter” lines**

Single copy Ds and Ac lines have been used as starter lines. From 1999, genetic crosses have been made to develop a large scale of transposant population. The following is a common genetic practice to obtain transposants via three generations in a single year. In
the first summer, F1 seeds are produced from the cross between Ac and Ds lines. In the first winter crop, F1 plants were selfed. In the second winter crop, transposants are selected from segregating F2 generation.

\[
P \quad \text{Ac X Ds} \quad \text{(summer field)} \\
\downarrow \\
F1 \quad \text{Ac/+ Ds/+} \quad \text{(1st Green-house (Sept – Jan)} \\
\downarrow \\
F2 \quad \text{screen for transposed Ds} \quad \text{(2nd Green-house)} \\
\quad \text{(Jan – May)} \\
\downarrow \\
F3 \quad \text{Field study of transposants} \quad \text{(the next summer field)}
\]

Based on Southern analysis on individual F2 progeny, germinal transmission rate has been determined in 2000 and 2001. In 2000, 25 percent or 18 percent of F2 plants of two lines, carried germinally transmitted transposed Ds. However, in 2001, the same combination of crosses showed lower rate of germinal transmission of transposed Ds. The frequency of transposed Ds in F2 was dropped to 12 percent in the line that showed 25 percent rate in the previous year. Average number of transposed Ds in F2 generation was 1.5 copies per plant. Since Ac and Ds stocks have been maintained by selfing, the loss of Ds transposition frequency might undergo inactivation of either Ac or Ds during self transmission. It is a common observation that many Ds element of F2 or later generations become inactive in plants expressing Ac mRNA.

b) Plant regeneration

Previous study implied that T1 transgenic plants carrying T-DNA harboring both Ac and Ds yielded much higher frequency of transposition than F2 from genetic crosses between Ac and Ds (Chin et al., 1999). Several independent transformations showed that over 70 percent of Ds were mobilized in calli before initiation of plantlets. From studies in maize and rice, it has been well demonstrated that plant regeneration from calli via tissue culture reactivates silent, inactive, or cryptic transposable elements (Peschke et al., 1987). To enhance the efficiency of generating Ds transposants, a large scale of tissue culture were performed to regenerate plants (R1) from calli homoyzogous for Ac and Ds. The number of Ds that were transposed before plantlet initiation, was much more than that of Ds that were transposed after plantlet initiation. The percentage of plants carrying Ds that was transposed during callus stage was 65 percent among total plants, and was 78 percent among the total Ds plants (Kim et al., 2002, Table 1). Average number of Ds transposed before plantlet initiation was 2.5 copies per plant.

<table>
<thead>
<tr>
<th>Total: 1224 regenerated plants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transposed Ds</td>
</tr>
<tr>
<td>Only original or somatic Ds</td>
</tr>
<tr>
<td>No Ds</td>
</tr>
</tbody>
</table>

Table 1. Summary of Southern analysis of regenerated plants from Ac/Ds seeds.

(Adapted from Kim et al., 2002)

c) Multi-copy Ds lines

Even though a single copy Ds line has been crossed to Ac lines, a small number (less than five percent) of F2 plants contained multi-copy (more than three copies) Ds. By examining the transmission of Ds in these multi-copy lines, the possibility was examined whether multi-copy Ds lines were maintained and propagated into a large population of transposant lines. Most Ds failed to be transmitted to the next generation. Furthermore, multi-copy Ds plants tended to show very low fertility and set very few seeds. We considered that generating multi-copy Ds lines might not be effective for a large number of transposant lines.

Limitation and consideration in the selection scheme

Since the mobility of Ac and Ds was clearly demonstrated, no large pool of Ds or Ac has been reported. Main reasons for delay in building a large pool of the transposants are lack of appropriate selection markers and spontaneous loss of activities of transposable elements. In maize and Arabidopsis, genetic markers have been commonly utilized to select transposons away from original insertion sites. For example, in Arabidopsis, negative dominant markers have been utilized to eliminate plants either carrying un-transposed elements or transposase ( Sudaresan et al., 1995). Excision markers have used to select elements transposed from original insertion sites. Our experience lead to the conclusion that selection schemes using genetic markers might not be effective in rice as much as in dicots such as Arabidopsis.
1) Inactivation of \( \text{Ac} \) and \( \text{Ds} \)

Our studies and others have demonstrated that \( \text{Ds} \) is highly active in early generations such as F1 or T1, however, in the subsequent generations, many \( \text{Ds} \) were quickly inactivated even in the presence of \( \text{Ac} \) (Izawa et al., 1997; Chin et al., 1999). Many \( \text{Ds} \) were inactivated even in the presence of \( \text{Ac} \) in plants of F2 or later generation and in plants regenerated from calli. Therefore, using a counter selection marker is not necessary to eliminate \( \text{Ac} \) and to stabilize transposed \( \text{Ds} \). Also, a well-known negative dominant selection system, cytochrome P450/R4702 (O’Keefe et al., 1994) was found not to be effective in rice (unpublished data). However, it has been argued that autonomous \( \text{Ac} \) itself might not be subjected to rapid inactivation in rice (Greco et al., 2001). It has been claimed that \( \text{Ac} \) can be maintained in many generations without loss of activity and, at the same time, increase the copy number. Therefore, it is necessary to check whether inactivation of the \( \text{Ac}/\text{Ds} \) family in rice genomes is a general phenomenon or not. Since many genetic factors related to chromatin structure and DNA methylation influence the mobility of transposable element, the difference could be due to genetic background (Miura et al., 2001; Singer et al., 2001).

2) Low efficiency of excision marker in selecting transposed \( \text{Ds} \)

In our tagging population, around 50 percent of the lines that contained transposed \( \text{Ds} \) (transposant lines) still carried original \( \text{Ds} \) (unpublished data). Therefore, a selection scheme using excision marker lead to loss of many \( \text{Ds} \) transposants.

Current strategy of \( \text{Ds} \)-tagging system and consortium effort of Korea

Based on the data from the three methods described above, the following strategy is proposed to rapidly generate a large scale of \( \text{Ds} \) transposants in rice. The initial distribution of \( \text{Ds} \) throughout the rice genomes is made with original \( \text{Ac} \) and \( \text{Ds} \) starter lines. By genetic crosses between \( \text{Ac} \) and \( \text{Ds} \) lines, F2 plants with both transposed \( \text{Ds} \) and untransposed \( \text{Ds} \) will be selected, based on chromosomal location. As mentioned above, most \( \text{Ds} \) of F2 or later generation are inactive even in the presence of \( \text{Ac} \) mRNA. We are selecting around 200 \( \text{Ac}/\text{Ds} \) starter lines that are evenly distributed on 24 arms of 12 chromosomes. The seeds of these lines will be subjected to tissue culture to regenerate plants. Among R1 plants, 65 percent of plants are expected to carry transposed \( \text{Ds} \). These \( \text{Ac}/\text{Ds} \) starter lines whose location of \( \text{Ds} \) is known, are suitable to perform targeted or local saturation mutagenesis.

A consortium for a \( \text{Ac}/\text{Ds} \) mediated Gene Tagging System has been established in Korea since 2001. Six labs throughout the country developed a consortium for this project. By co-operative efforts, \( \text{Ds} \) transposants are generated by genetic crosses and regeneration. To regenerate transposant lines via tissue culture, \( \text{Ac}/\text{Ds} \) starters are selected and propagated during summer. During winter, around 20 000 regenerated plants are produced. Since 65 percent of the population carry transposed \( \text{Ds} \), DNA is prepared from a pool of two plants for TAIL-PCR to clone \( \text{Ds} \) flanking DNA sequence. Every year, 10 000 \( \text{Ds} \) lines will be selected. 4500 \( \text{Ds} \) loci are cloned and sequenced. Based on this data, the consortium plans to develop a database for tagged genes.

Characterization of \( \text{Ds} \) insertion sites

1) GUS expression mediated by gene trap \( \text{Ds} \)

The probability of \( \text{Ds} \) insertion into a gene has been estimated by enhancer trap \( \text{Ds} \) in Arabidopsis (Martienssen, 1998). It appeared true that the majority (up to 80 percent) of \( \text{Ds} \) insertion sites is at or near a genic locus. In Arabidopsis, 26 percent transposants carrying gene trap \( \text{Ds} \) exhibited GUS stains in seedlings, leaves, or flowers (Sundaresan et al., 1995). About eight percent total transposed \( \text{Ds} \) showed GUS stains in floral organs of rice (Chin et al., 1999). Therefore, the efficiency of gene identification via GUS stains should be comparable between rice and Arabidopsis. Since the expression of GUS is under the control of transcriptional machinery of a host gene, genes responding to environmental or developmental signals can be examined in plants.

2) Cloning of \( \text{Ds} \) flanking genomic DNA

Reverse genetics is advantageous once a large scale of insertional mutagenesis was performed. However, it has to go through the same DNA samples via PCR every time when a new tagged gene is sought from a population. Many facilities of functional genomics have collected the sequence information of flanking DNA of either transposable elements or T-DNA and have developed databases of flanking DNA sequences, which makes the analysis of mutagenesis population much simpler and easier. Since the insertion mode of T-DNA is more complicated than that of transposable elements, transposable elements are a much better choice in
obtaining flanking genomic DNA (Tax and Vernon, 2001).

In our group, two PCR techniques, namely IPCR (inverse PCR) and TAIL-PCR (thermal asymmetrical interlaced PCR), have been applied and compared to clone genomic DNA flanked with Ds. More than 80 percent of Ds flanking DNA were successfully amplified by iPCR. For TAIL-PCR, choice of appropriate arbitrary degenerate (AD) primers greatly enhanced the success ratio of TAIL-PCR. Four AD primers among a set that Dr. Don McCarthy developed for cloning Mu-flanking DNA of maize genome, produced a high success rate. More than 90 percents of Ds elements produced PCR products by applying two AD primers to a given population. Compared to TAIL-PCR, iPCR uses specific primers and, therefore, achieve a high success rate of cloning target DNA.

However, it required many molecular manipulation steps before PCR reaction. Furthermore, if initial digestion is not complete due to either partial DNA modification or impure DNA samples, iPCR frequently failed to yield products. TAIL-PCR requires very simple manipulation before PCR amplification, therefore, it is suitable for mass production of Ds flanking DNA once appropriate degenerate primers and strenuous Taq polymerase are employed in PCR reaction.

The data clearly demonstrated that, Ds tagging system should generate excellent genetic materials and molecular information in rice.

### 3) Spectrum of the putative functions of genes inserted by Ds

Up to now (20 April 2002), more than 1000 Ds flanking DNA have been sequenced. Among them, 429 Ds-insertion sites have been analyzed, via database searches, to define a coding region of putative function. 12 clones (2.8 percent) carried T-DNA vector sequences. 10.5 percent are clones from identical Ds. Among the rest 372 genes, 207 (48.3 percent) showing at least 50 percent homology with functional domains or coding regions of previously reported proteins. There is a spectrum of functional diversity in Ds-inserted rice genes. Many genes encoding kinases, receptors, transcription factors, and signal mediating molecules such as small G proteins and Ring zinc finger protein, have been tagged and “knocked-out” by Ds. Even though the number of genes inserted by Ds is still small, the data clearly demonstrated that, Ds tagging system should generate excellent genetic materials and molecular information in rice.

<table>
<thead>
<tr>
<th>T-DNA region</th>
<th>12</th>
<th>2.8 percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duplicate clones from identical Ds</td>
<td>45</td>
<td>10.5 percent</td>
</tr>
<tr>
<td>No or low homology</td>
<td>165</td>
<td>38.5 percent</td>
</tr>
<tr>
<td>Significant homology (&gt; 50 percent)</td>
<td>207</td>
<td>48.3 percent</td>
</tr>
<tr>
<td>Total</td>
<td>429</td>
<td>100 percent</td>
</tr>
</tbody>
</table>

### 4) Mapping of transposed Ds

It is well documented that many Ds are relocated around their original site. To examine the distribution of Ds on rice chromosomes, transposed Ds loci originated from two different Ds starter lines were cloned and sequenced. One line showed that 24 percent transposed Ds were at T-vector sequence while the other line gave only three percent T-DNA sequence. Therefore, the mode of transposition appeared to depend either on the location of Ds in rice genomes or on subtle structural differences of original Ds sites that could be expected during integration of T-DNA. From mapping information obtained either using recombinant inbred lines or by database searches indicated that we did not find any distinct preferential distribution of transposed Ds. Ds is more or less relocated throughout all 12 chromosomes. These mapped Ds will be used for targeted or local saturation mutagenesis.

### Further consideration in operating knock-out systems

#### 1) Gene duplication and functional redundancy

Visual expression at the phenotypic level is very crucial in determining the agronomical or biological significance of any given gene. However, multiplication of genes of the same protein family complicates the analysis of gene functions and their phenotypes. To detect a phenotype of a mutant, it might be required to generate double or triple or even more combinations of mutations (Martienssen and Irich, 1997). Under this circumstance, the real function of a single gene is very difficult to be evaluated. To reveal “real” function of a
gene determining phenotypic character, recessive null mutation is not sufficient. It requires ectopic expression or RNAi expression of genes of interest. Activation tagging population or systemic generation of RNAi population should be prepared as a complementary genetic tool for functional genomics in rice.

2) Genetic Background

Most rice cultivars have arisen from three subspecies of genetic source. These are japonica, indica, and javanica. Many agronomical characters of these subspecies are different. Since draft sequences have been published from the japonica and indica varieties, many genetic differences can be expected and detected based on the sequence difference. Due to low transformation efficiency of indica plants, most gene tagging systems have been established using japonica rice. Traditionally, japonica varieties have been developed to improve grain quality while indica varieties have been focused on increasing characters more directly improving yield, such as disease and plant shape. It is not practical to establish a large scale of T-DNA insertional mutants in the indica genetic background, as the transformation efficiency is low. However, transposable element can be utilized since not many starter lines are required to mutagenize indica genomes. International efforts are required to develop gene tagging systems in indica genetic backgrounds.

Acknowledgments

I am very grateful to all the members of my lab who provided unpublished data for this review and Dr. Eu Moo Yong, the head of the Ac/Ds gene tagging project, for his support. This work was supported by a grant (code #171 & 1713) from Crop Functional Genomics Center of the 21st Century Frontier Research Program funded by the Ministry of Science and Technology, Republic of Korea and by a grant No.R03-2001-00072 from the Basic Research Program of the Korea Science and Engineering Foundation.

References


