The Plant Journal (2005) **44**, 879–892 doi: 10.1111/j.1365-313X.2005.02570.x

TECHNICAL ADVANCE

Efficient insertional mutagenesis in rice using the maize *En/Spm* elements

Chellian Santhosh Kumar[1], Rod A. Wing[2] and Venkatesan Sundaresan[1,3,*]
[1]*Department of Plant Sciences, Life Sciences Addition 1002, University of California, One Shields Ave, Davis, CA 95616, USA,*
[2]*Arizona Genomics Institute, 303 Forbes Building, University of Arizona, Tucson, AZ 85721, USA, and*
[3]*Section of Plant Biology, Life Sciences Addition 1002, University of California, One Shields Ave, Davis, CA 95616 USA*

Received 20 April 2005; revised 31 August 2005; accepted 2 September 2005.
*For correspondence (fax +1 530 752 5410; e-mail sundar@ucdavis.edu).

Summary

We have developed a novel system for insertional mutagenesis in rice (*Oryza sativa*) based on the maize (*Zea mays*) enhancer/suppressor mutator (*En/Spm*) element. In this system, a single T-DNA construct with *Spm*-transposase and the non-autonomous defective suppressor mutator (*dSpm*) element is used in conjunction with green fluorescent protein (GFP) and *Discosoma sp.* Red Fluorescence Protein (DsRed) fluorescent markers to select unlinked stable transpositions of *dSpm*. Using this system, we could demonstrate high frequencies of unlinked germinal transposition of *dSpm* in rice. Analysis of *dSpm* flanking sequences from 353 stable insertion lines revealed that the *dSpm* insertions appear to be widely distributed on rice chromosomes with a preference for genic regions (70%). The *dSpm* insertions appear to differ from Activator-Dissociation (*Ac-Ds*) elements in genomic distribution and exhibit a greater fraction of unlinked transpositions when compared with *Ds* elements. The results obtained in this study demonstrate that the maize *En/Spm* element can be used as an effective tool for functional genomics in rice and can complement efforts using other insertional mutagens. Further, the efficacy of the non-invasive fluorescence-based selection system is promising for its application to other crops.

Keywords: *Oryza sativa*, *En/Spm-I/dSpm*, transposon tagging, unlinked transposition, *dSpm* flanking sequence tags, functional genomics.

Introduction

Rice (*Oryza sativa*) is one of the most important crops in the world (Khush, 1997) and has important syntenic relationships with other cereal species such as sorghum (*Sorghum bicolor*), maize (*Zea mays*), barley (*Hordeum vulgare*) and wheat (*Triticum aestivum*) (Delseny, 2004; Gale and Devos, 1998). As a result of its small genome size and its long history as a subject of genetic and molecular studies it has become a model for cereal crops. The draft genome sequence for two subspecies of rice, *Oryza sativa* ssp. *japonica* (Goff *et al.*, 2002) and *Oryza sativa* ssp. *indica* (Yu *et al.*, 2002), has already been reported. The complete high-quality sequencing of the *japonica* rice cultivar Nipponbare is near completion by the International Rice Genome Sequencing Project (http://rgp.dna.affrc.go.jp/IRGSP/). Gene predictions in The Institute for Genomic Research (TIGR) database have identified 57 915 genes (loci) from 370 630 783 bp of non-overlapping rice genome sequence (http://www.tigr.org/tdb/e2k1/osa1/pseudomolecules/info.shtml). The availability of the complete genome sequence emphasizes the need for tools that will permit assignment of functions to these predicted genes, which are now known by sequence alone. Insertional mutagenesis is one of the most suitable ways to disrupt genes and understand their function. Furthermore, the identification of functions of rice genes will provide valuable information for investigating similar traits or regions in the genomes of other grass species (Bennetzen and Ma, 2003; Gale and Devos, 1998).

Insertional mutagenesis has been successfully carried out in rice using T-DNA (Jeong *et al.*, 2002; Sallaud *et al.*, 2004; Wu *et al.*, 2003) and the endogenous retrotransposon Transposon *Oryziva sativa* 17 (*Tos17*) (Hirochika, 2001; Hirochika *et al.*, 1996; Miyao *et al.*, 2003). Several laboratories have also developed strategies using the heterologous transposon Activator (*Ac*) and the Activator/Dissociation (*Ac/Ds*) system to generate insertion lines (Chin *et al.*, 1999; Enoki *et al.*, 1999; Greco *et al.*, 2001, 2003; Ito *et al.*, 2004; Izawa *et al.*, 1997; Kim *et al.*, 2004; Kohli *et al.*, 2001; Kolesnik *et al.*, 2004; Nakagawa *et al.*, 2000; Upadhyaya *et al.*, 2002).

The autonomous enhancer/suppressor mutator (*En/Spm*) transposable element was originally identified in maize by Peterson (1953) and McClintock (1954) and molecularly characterized by Pereira *et al.* (1986) and Masson *et al.* (1987). This family of transposable elements in maize consists of an autonomous transposition competent *En/Spm* element and a non-autonomous inhibitor/defective *Spm* (*I/dSpm*) element. The complete 8.3-kb autonomous *En/Spm* element contains the transposase gene and promotes its own transposition, as well as that of *I/dSpm*. *I/dSpm* elements have the same terminal sequences as *En/Spm* but carry internal deletions, which render them incapable of encoding functional transposase. Members of the *En/Spm* transposable element family have a 13-bp terminal inverted repeat and create a 3-bp target site duplication on insertion.

In order to develop a new efficient system for large-scale mutagenesis in rice, we chose to investigate the maize *En*/*Spm* transposon family, which is known to work well in the heterologous host Arabidopsis (Aarts *et al.*, 1995; Marsch-Martinez *et al.*, 2002; Speulman *et al.*, 1999; Tissier *et al.*, 1999; Wisman *et al.*, 1998). Studies in Arabidopsis have shown that *En/Spm* elements transpose with high efficiency (Aarts *et al.*, 1995).

In a previous study, the *En/Spm* system was introduced into rice but low frequencies of somatic transposition were obtained and germinal transposition in rice could not be clearly demonstrated (Greco *et al.*, 2004). Here we describe a novel fluorescence-based tagging strategy based on the *En/Spm* transposon, which enables easy recovery of stable unlinked single-copy *I/dSpm* element insertions in rice. Using our strategy, we were able to efficiently generate stable new insertions of *dSpm*, demonstrating its feasibility for large-scale insertional mutagenesis in the rice genome.

Results

Constructs and selection of stable transposants

The T-DNA vector carrying the immobilized *Spm* and defective *Spm* (*dSpm*) elements *in cis* is shown in Figure 1(a). The immobilized *Spm*-transposase source was derived by deleting the ends of the complete autonomous *Spm* transposon, so that it could no longer transpose. The Cauliflower Mosaic Virus (CaMV) 35S promoter was used to drive the expression of *Spm*-transposase transcript, which encodes the proteins required for the transposition of the *dSpm* element in the rice genome. The *dSpm* element was generated by an internal deletion of 6.2 kb from the complete *Spm* element, retaining 1014 bp at the 5′ terminus and 1193 bp at the 3′ terminus. It carries a positive selection marker, the *Discosoma sp.* Red Fluorescence Protein

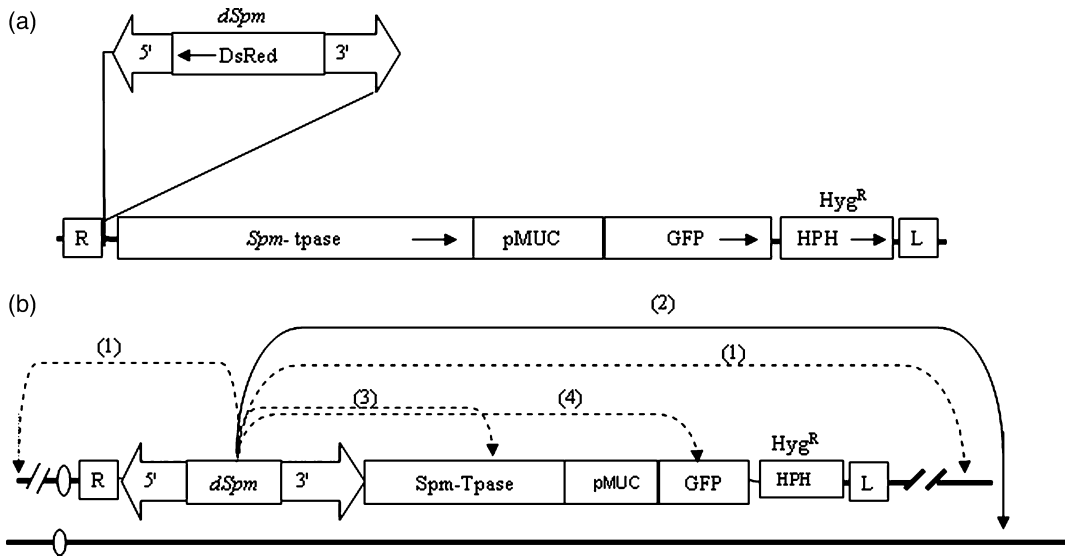

Figure 1. (a) T-DNA region of the construct used in this study. (b) Possible transposition events for the defective suppressor mutator–*Discosoma sp.* Red Fluorescence Protein (*dSpm*-DsRed) element. (1) Unlinked or loosely linked transposition to same chromosome; (2) unlinked transposition to a different chromosome; (3) closely linked transposition into the *Spm*-transposase gene; (4) closely linked transposition into the green fluorescent protein (*GFP*) gene. R and L, right and left borders of the T-DNA, respectively; GFP, negative selection marker; DsRed, red fluorescence protein gene; HPH, hygromycin resistance gene selection marker for transformation.

(*DsRed*) gene, which encodes a red fluorescence protein (Clontech, Mountain View, CA, USA) resulting in red fluorescence with an emission maximum at 583 nm when excited with 558 nm. The *DsRed* gene is driven by the maize ubiquitin 1 promoter (Christensen and Quail, 1996) and has a nopaline synthase (Nos) termination signal at the 3′ end. The element was designed in such a way that the Nos termination sequence and the termination sequence of the *dSpm* element are in opposite orientations. The presence of two terminators helps to ensure that insertion of the element in either orientation will disrupt the gene function even if the element inserts into an intron. To select for unlinked transposition events and to select against the *Spm*-transposase, we used the synthetic green fluorescent protein (sGFP) as the negative selection marker (Chiu *et al.*, 1996). The sGFP has an excitation maximum at 488 nm and an emission maximum at 507 nm and is driven by the maize ubiquitin1 promoter. The vector also carries the CaMV 35S promoter-driven Hygromycin Phosphotransferase (*HPH*) gene conferring resistance to the antibiotic hygromycin B, which was used to select transformants. We chose a combination of fluorescent protein marker genes (GFP and DsRed) as negative and positive selection markers, respectively, to enable quick and easy screening of the germinated seedlings. The other major advantage of using GFP is that it makes it feasible to distinguish between homozygous and heterozygous plants based on GFP intensity.

The principle of selecting unlinked transposition events is represented in Figure 2(a). In the transgenic plants carrying the complete T-DNA, we expect that transposition of the *dSpm* element will occur to both linked and unlinked sites in the genome. The different types of possible transposition events are illustrated in Figure 1(b). In plants with a single T-DNA locus, 25% of the selfed progeny will be GFP^-, as the GFP is not mobile. If the *dSpm* element carrying DsRed transposes to a site that is unlinked or loosely linked to the donor site, then a fraction of the progeny that are GFP^- will be $DsRed^+$. These progeny can be selected and propagated as stable insertion lines (stable transposants), as they do not carry the *Spm*-transposase gene. GFP^-;$DsRed^+$ progeny can also result from closely linked transposition of the *dSpm* element into the *GFP* gene itself, resulting in some lines that contain insertions within the T-DNA rather than in the rice genome. In our study, such lines could be identified based on somatic reversion sectors (GFP-positive sectors) resulting from *dSpm* excision or by polymerase chain reaction (PCR) using primers specific to the T-DNA regions or from the flanking sequence data (see below).

Generation of starter lines and detection of dSpm *transposition in rice*

The immobilized *Spm–dSpm* T-DNA construct was introduced into *Oryza sativa* ssp. cv. Nipponbare by *Agrobacterium tumefaciens*-mediated transformation using scutellum-derived calli from mature seeds (see Experimental procedures). The excision of the *dSpm* element from the original T-DNA position was monitored by PCR using DNA isolated from 28 primary transformants (T_1 transformants). The results indicated that 19 out of 28 (67.8%) double fluorescent (GFP^+ and $DsRed^+$) T_1 transformants contained the empty donor sites (EDS) which could have arisen from somatic or germinal transposition. Of the 19 that contained EDS, 16 contained both the original full donor site (FDS) and the empty donor site (EDS) as expected of somatic transpositions. The *dSpm* element used in our study is derived from *Spm* insertion into the maize *A1* gene (Masson and Fedoroff, 1989) which has an additional 3-bp insertion adjacent to one end of the element. The sequence analysis of the EDS fragments from 14 transformants revealed the presence of molecular footprints left behind after the *dSpm* excision. Of the 14 EDS sequences, 12 had precise excision of the *dSpm* element and two had a 1-bp insertion at the same position (Table 1).

Selection of active lines for pilot-scale insertion line production

For our strategy to generate insertion lines, we needed to identify heterozygous T_2 seedlings from transformants with single-copy or single-locus T-DNA (Figure 2a). This was done for all the 28 GFP- and DsRed-expressing transformants based on the segregation of GFP fluorescence. Of the 28 transformants, 22 (78.5%) segregated for GFP in a 3:1 ratio in the T_2 generation. Of these 22 single-locus transformants, four had very low levels of GFP expression which could interfere with the prediction of zygosity (see below) and two showed low fertility. These lines were eliminated and the remaining 16 transformants were selected for further production of insertion lines.

In order to ensure as many independent events as possible, we needed to avoid propagating transpositions that occurred early in the primary transformant (T_1 plant) resulting in multiple T_3 families with the same insertion. Recovery of such insertions can be minimized by not using seeds from any T_1 plants showing a significant number of GFP^- $DsRed^+$ T_2 progeny (two of the 28 T_1 plants were in this category).

Heterozygous T_2 seedlings from these 16 selected transformants were identified based on the intensity of GFP fluorescence levels (Figure 2b). These were planted in the glasshouse and self-pollinated to obtain T_3 seeds, which were used to screen for stable transposants. T_3 seeds from 861 putative T_2 heterozygous plants were obtained and GFP segregation in the T_3 progeny confirmed that 803 (93.26%) of the T_2 plants were heterozygous for the T-DNA and the rest were homozygous (Table 2). The accuracy of selecting heterozygous T_2

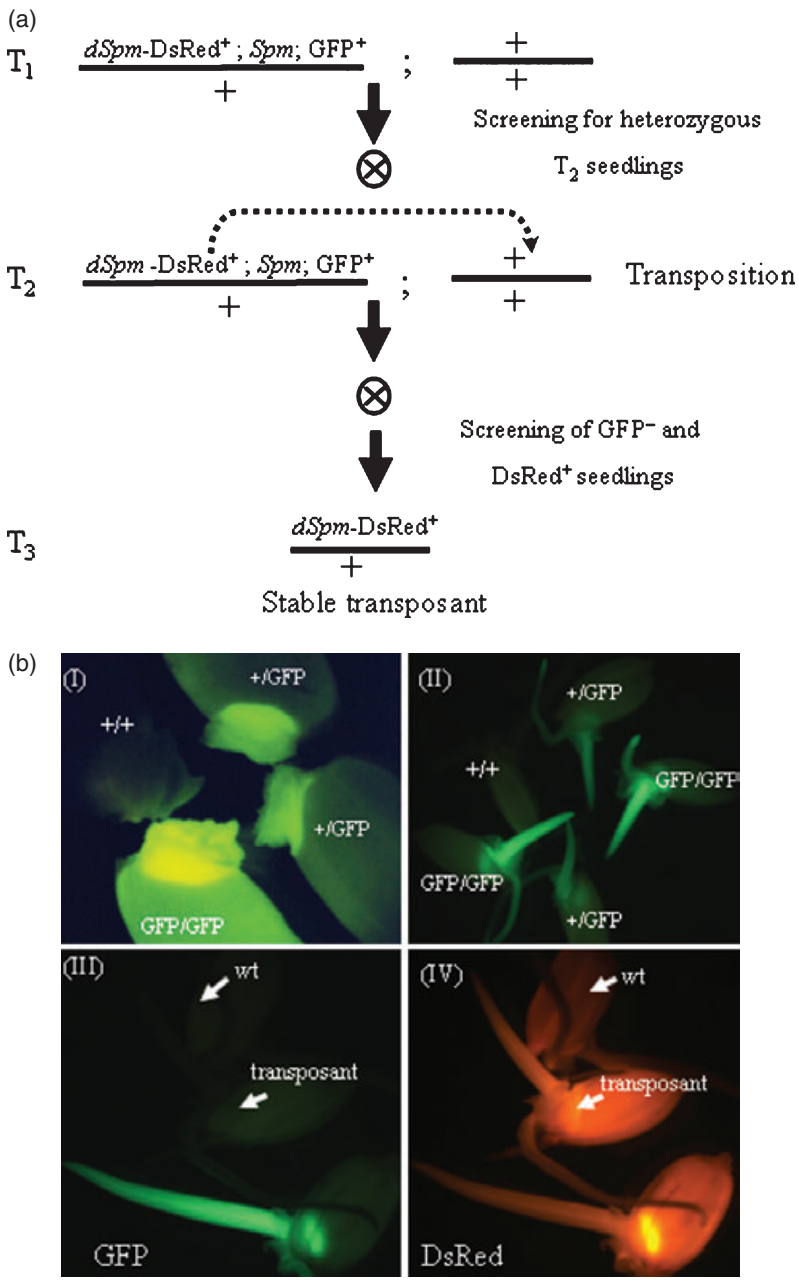

Figure 2. (a) *In cis* strategy for generating insertions. (b) (I and II) Segregation of green fluorescent protein (GFP) (1:2:1) in progeny of a single-locus T_1. (III and IV) Screening for stable transposants by GFP and *Discosoma sp.* Red Fluorescence Protein (DsRed) fluorescence. White arrows indicate a putative transposant which is GFP^- (left panel) and $DsRed^+$ (right panel).

Table 1 Empty donor site footprints from defective suppressor mutator (*dSpm*) excision in T_1 plants

	Sequence	Frequency of EDS (number of T_1 plants analyzed)
Original FDS	ATCCTCCATT *dSpm* AATATTCAA	
EDS 1	ATCCTCCATT AATATTCAA	12
EDS 2	ATCCTCCATT **A**AATATTCAA	1
EDS 3	ATCCTCCATT AAT**T**ATTCAA	1

FDS, full donor site; EDS, empty donor site.

seedlings by the intensity of GFP varied among the different transformants (80 to 98%). The pre-selection of heterozygous T_2 seedlings using GFP intensity helps to reduce the number of T_2 homozygotes, which cannot be used to generate stable transposants because the *Spm*-transposase cannot be segregated.

Selection of stable transposants

Screening for stable transposants is performed in the T_3 generation. Approximately 100–250 seeds from each T_2 plant were germinated in water. After 4–5 days these seed-

Table 2 Estimation of germinal unlinked transposition frequency

Transformant ID	Putative T_2 hets planted	Confirmed T_2 hets	T_3 families with ≥1 transposant (GFP^-DsRed^+)	Putative unlinked germinal TF (%)
1	80	77	43	55.8
2	48	45	11	24.4
3	75	72	8	11.1
5	32	31	0	0
6	26	25	12	48
7	50	49	14	25
8	56	54	12	22.2
12	50	42	8	19.0
16	20	19	3	15.8
20	74	70	17	24.3
21	50	40	34	85
22	60	57	36	63.2
24	50	48	9	18.7
25	60	57	41	71.9
28	60	55	43	78.2
29	70	62	46	74.2
Total	861	803	337	41.96

The putative germinal unlinked transposition frequency was calculated as the percentage of GFP^- and $DsRed^+$ T_3 families out of the total number of T_3 families screened.
hets, heterozygotes; TF, Transposition Frequency; GFP, green fluorescent protein; DsRed, *Discosoma sp.* Red Fluorescence Protein.

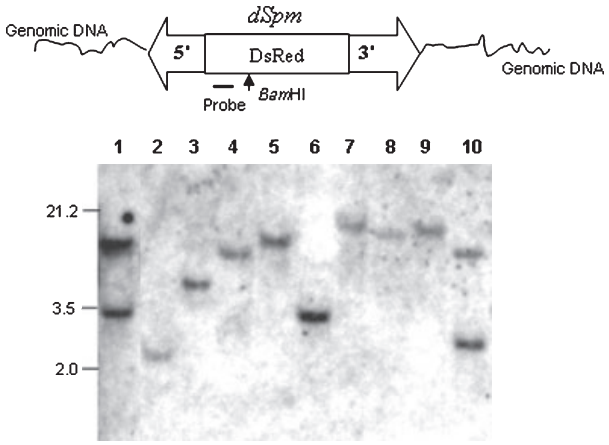

Figure 3. Southern hybridization analysis on putative stable insertion lines. Genomic DNA from T_3 transposants was digested with *Bam*HI, and the blots were probed with the *Discosoma sp.* Red Fluorescence Protein (*DsRed*) gene. The transposed defective suppressor mutator (*dSpm*) element should give a band of more than 2 kb. Lane 1, T_1 parental line; lanes 2–10, putative stable transposants.

lings were screened for GFP^- and $DsRed^+$ phenotype. Of the 803 heterozygous T_3 families screened, 337 (41.96%) gave at least one transposant (i.e. GFP^-;$DsRed^+$) seedling. The results are summarized in Table 2. The frequency of germinal unlinked transposition varied from 0 to 85% among the different transformants. In three T_3 families the $DsRed^+$ GFP^- seedlings had revertant sectors of GFP expression, indicating the presence of *dSpm* insertions in the *GFP* gene. All of these originated from a single transformant (number 6). Transformant 5, which did not produce any stable transposants (none of 31) in the T_3 generation, was also negative for the presence of an EDS in the T_1 generation. However another transformant, number 16, which did not show an EDS in the T_1 generation, did produce stable transposants at low frequency (three of 19) in the T_3 generation.

Randomly selected GFP^-;$DsRed^+$ seedlings, representing putative stable transposants, were confirmed for the presence of *dSpm* and the absence of a donor site by PCR reaction using primers specific to the T-DNA and *dSpm*-DsRed. The *HPH* gene-specific primers were used to determine whether the putative transposants contained the original T-DNA. The results indicated that 35 of 36 plants tested were negative for the *HPH* gene, indicating that the GFP^- screening was effective in selection for the absence of a donor site. The one plant out of 36 that still had the T-DNA probably corresponds to a re-insertion within the *GFP* gene, as shown in Figure 1(b). The PCR test for the *DsRed* gene showed that all the selected $DsRed^+$ plants based on red fluorescence contained the *DsRed* gene, confirming that there were no false positives in the DsRed selection. The insertion lines were further characterized by genomic Southern blot hybridization using the *DsRed* gene as the probe. Southern blot hybridization on putative stable transposants revealed the presence of independent insertions of the *dSpm* element based on the different sizes of the border fragments (Figure 3). In nine out of 10 transposants, there was single copy of the transposed *dSpm* element, consistent with the 3:1 genetic ratios observed for *DsRed* marker segregation among the transposants (data not shown).

To summarize, a total of 28 starter lines were generated, of which 22 were single locus based on segregation of T-DNA. Of these single-locus starter lines, 16 exhibited good expression of GFP and DsRed and high fertility. Of the 16 starter lines, 11 yielded most of the insertions used for flanking sequence tags (FST) analysis (Table 2). These 11 starter lines were further molecularly characterized by Southern hybridization and flanking sequence analysis. Southern hybridization revealed that six of the starter lines had single-copy T-DNA (numbers 1, 3, 8, 12, 21 and 25), three of them contained two copies (numbers 2, 7 and 20) and two had multiple copies (numbers 6 and 24). Flanking sequences of 10 of these 11 starter lines were obtained and used to map the position of the integrated T-DNA locus on rice chromosomes (Table 3).

Analysis of En/Spm *transcripts in rice*

The active transposases required for the transposition of the maize *En/Spm* element have been shown to originate

Table 3 Location of defective suppressor mutator (*dSpm*) T-DNA in starter lines on physical maps

Transformant ID	Chr	BACs/PACs	Insertion position (bp)	Genetic map position (cM)
1	1	P0435B05	95704	129.3
2	ND	–	–	–
3	9	P0681H07	92805	2.1
6	8	OSJNBa0012O03	24565	57.8
7	4	OSJNBa0010H02	93042	85.5
8	3	OSJNBa0066H15	26558	96.6
12	2	P0527E02	144018	56.8
20	10	OSJNBa0027L23	110974	83
21	1	P0494A10	51455	5.6
22	1	P0443E05	151534	32.4
25	7	P0047B07	80068	103.4

ND, not determined; Chr, chromosome; BAC, Bacterial Artificial Chromosomes; PAC, PI-derived artificial chromosomes.

from alternate splicing of the single large *En/Spm* transcript (Masson *et al.*, 1989; Pereira *et al.*, 1986). We assayed for the presence of such alternatively spliced transcripts by reverse transcriptase–polymerase chain reaction (RT-PCR) in a few selected lines using primers designed to anneal at positions along the maize *Spm* element, similar to Masson *et al.* (1989). A total of three combinations of primers were used to check for the presence of transposase A (TnpA), B (TnpB), C (TnpC) and D (TnpD) as in Greco *et al.* (2004). Analysis of the RT-PCR products obtained revealed alternate splicing of the *Spm* transcript, as has been shown previously in maize (Masson *et al.*, 1989; Pereira *et al.*, 1986) and in rice (Greco *et al.*, 2004). Of the transformant lines analyzed, transformants 21 and 25 seemed to have higher overall Tnp transcript levels (Figure 4b), which might account for the higher frequencies of unlinked transposition in the T_3 generation (85 and 71.9%, respectively). In transformant 3, which gave a very low frequency of unlinked transposition (11.1%), no RT-PCR products were detected, which might indicate very low levels of transcripts. Similarly, in transformant 5, which did not produce any stable transposants, we did not find any RT-PCR products, which could be attributable to low levels or complete absence of the transcript. Our RT-PCR results indicated that the transcript TnpD is more abundant than TnpA, TnpB and TnpC. To further determine the relative abundances of TnpD and TnpA, which are required for transposition, we carried out semiquantitative RT-PCR. After 20 cycles of amplification the TnpD product was visible but not the TnpA product, reflecting the greater abundance of the TnpD transcript relative to the TnpA transcript (Figure 4c). The TnpA transcript was visible after 25 cycles of PCR.

Analysis of dSpm *flanking sequence tags (FSTs)*

The *dSpm* flanking sequences from different putative stable transposants (GFP^- and $DsRed^+$) were generated by Thermal-Asymmetric Interlaced-PCR (TAIL-PCR). A total of

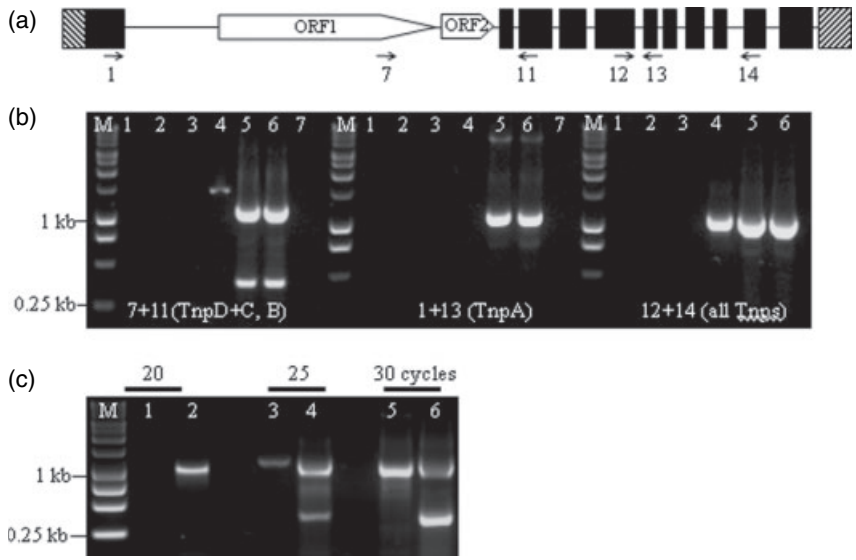

Figure 4. Analysis of enhancer/suppressor mutator (*En/Spm*) transcripts in transgenic rice plants.
(a) Sketch of the maize *En/Spm* element showing the locations of primers used for reverse transcriptase–polymerase chain reaction (RT-PCR). Subterminal repetitive regions, exons and the open reading frames (ORFs) in the first introns are represented as hatched boxes, black boxes and open arrows, respectively.
(b) A Gel Star stained agarose gel showing RT-PCR products obtained with different primer combinations on the *En/Spm* transcript. Primer combination 7 + 11 amplified 1.28 kb of Transposase D (TnpD), 1.18 kb of TnpC and 0.49 kb of TnpB transcripts. Primer combination 1 + 13 amplified a 1.1-kb region of TnpA. Primer combination 12 + 14 amplified a 0.97-kb region of all transcripts. M, a 1-kb DNA ladder (Promega, San Luis Obispo, CA, USA). Lane 1, rice wild type; lane 2, stable defective suppressor mutator (*dSpm*) insertion line; lane 3, Transformant (TF) 5; lane 4, TF 3; lane 5, TF 21; lane 6, TF 25; lane 7, water control.
(c) A Gel Star stained agarose gel showing semiquantitative RT-PCR products. Lanes 1, 3 and 5, TnpA (primers 1 + 13); lanes 2, 4 and 6, TnpD (primers 7 + 11).

353 unique FSTs were obtained. These transposants were derived from 11 independent primary transformants. The different sequences obtained were compared to public databases for similarity to known genomic sequences, Expressed Sequence Tags (ESTs) and protein sequences. Of the 353 unique FSTs, 341 (96.6%) represented insertions in the rice genome and 12 (3.4%) represented insertions into the T-DNA. Of the 12 insertions within the T-DNA, five originated from a single transformant (number 6), suggesting that this transformant might be atypical, and the remaining seven originated from six different transformants. The validity of the recovered flanking sequences was verified in 20 randomly selected insertion lines by PCR using primers specific to the flanking sequence and the *dSpm* ends. The results indicated that 20 of 20 lines showed expected products, using primers for both ends (18 of 20) or for the *dSpm* 3′ end (two of 20). All the flanking sequences obtained were submitted to GenBank and can be accessed publicly at http://sundarlab.ucdavis.edu/rice/blast/blast.html.

Chromosomal distribution of new dSpm *insertions*

The physically mapped *dSpm* insertions were linked to the genetic map *in silico* using the data from the TIGR and Rice Genome Research Program (RGP) databases (Figure 5). The pattern of transposition was analyzed for 312 insertion lines originating from nine transformants in relation to the original T-DNA start position (Table 4). Overall, 1.92% (six of 312) of the new *dSpm* insertions were within 4 cM of the original T-DNA position; these transpositions originated from three of nine T_1 transformants analyzed. When analyzed for intrachromosomal versus interchromosomal insertions, 12.17% (38 of 312) of the insertions were found to be intrachromosomal and the remaining 87.8% (274 of 312) were interchromosomal.

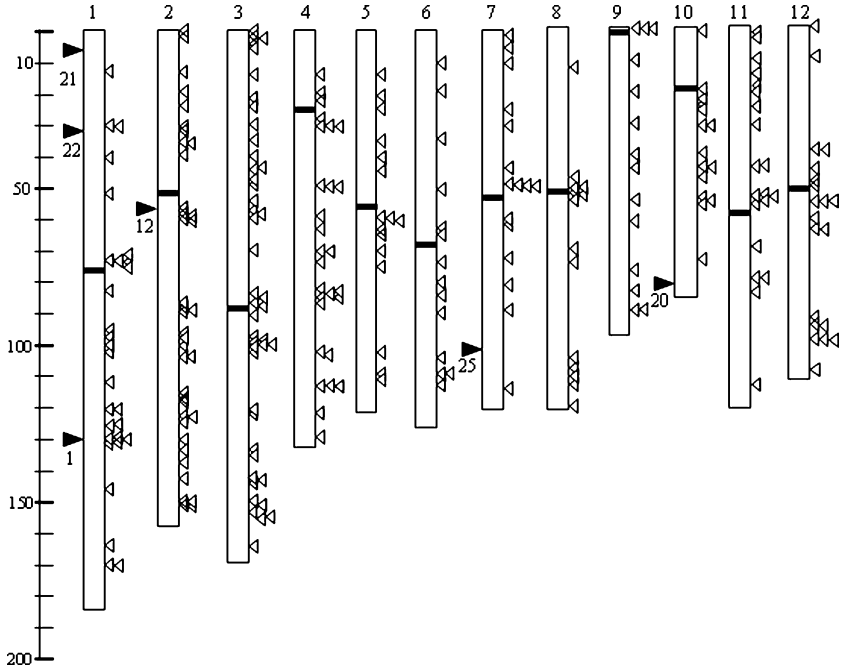

Figure 5. Distribution of genetically mapped defective suppressor mutator (*dSpm*) insertions on rice chromosomes. Horizontal black lines on the chromosomes are the centromere, black triangles are the original *dSpm* positions and open triangles mark the positions of new *dSpm* insertions.

Table 4 Distribution of new defective suppressor mutator (*dSpm*) insertions from the original T-DNA locus

Transformant ID	Insertion within 4 cM	Intrachromosomal insertions	Interchromosomal insertions	Total number of insertions
1	3	7	51	58
3	0	0	10	10
7	0	0	17	17
8	1	6	10	16
12	2	8	19	27
20	0	1	23	24
21	0	6	56	62
22	0	5	20	25
25	0	5	68	73
Total	6	38	274	312

Table 5 Examples of defective suppressor mutator (*dSpm*) insertions within annotated rice genes (a full list of insertions is available in Table S1)

Line	Gene ID	Putative function (TIGR rice genome annotation)
303	Os01g71310	FAD-binding domain, putative
310	Os03g51690	Homeobox 1 protein OSH1
342B	Os01g01840	Helix-loop-helix DNA-binding domain, putative
351	Os07g20580	Helicase conserved C-terminal domain
354A	Os01g55750	TCP family transcription factor, putative
355	Os02g46260	Serine carboxypeptidase
356	Os10g25560	Myosin head (motor domain), putative
358B	Os05g46370	Helix-loop-helix DNA-binding domain, putative
370A	Os02g42110	Protein kinase domain, putative
405	Os03g56020	Putative RNA-binding protein
407	Os01g66760	Leucine-rich repeat
411A	Os08g41720	Auxin efflux carrier
413B	11218.m00205	Cytochrome P450 reductase
420C	Os04g28130	myb-like DNA-binding domain, SHAQKYF class, putative
422	Os02g01090	Similar to transcriptional regulator superman
422C	Os03g51110	Putative Myb-like DNA-binding protein
427D	Os06g50150	Calcium-dependent protein kinase CPK1
436B	Os06g51410	Prolyl oligopeptidase family
439C	Os08g44230	DHHC zinc finger domain
445A	Os03g61740	Putative short chain dehydrogenase/reductase
455	Os02g06160	Protein kinase domain, putative
458	Os05g04640	WRKY DNA-binding domain, putative
465	Os06g36330	MATE efflux family protein, putative
467A	Os10g20880	Protease inhibitor/seed storage/LTP family, putative
476A	Os09g03930	ABC transporter, putative
482	Os03g17480	Putative glutathione S-transferase
483C	10324.m00143	Transmembrane amino acid transporter protein
490A	Os03g56020	Putative RNA-binding protein
494	Os04g55660	GDSL-like lipase/acylhydrolase
495	Os04g46450	Zinc finger, C3HC4 type (RING finger), putative

ABC transporter, ATP-binding Cassette transporter; GDSL, motif with Gly, Asp, Ser, and Leu; FAD, Flavin Adenine Dinucleotide; MATE, multidrug and toxin extrusion; RING, Really Interesting New Gene; TCP, TB1, CYC and PCF.

dSpm *insertions are preferentially in genic regions*

The flanking sequences were analyzed to determine if there was a preference for insertions into genic regions. The TIGR database (http://www.tigr.org/tdb/e2k1/osa1/pseudomolecules/info.shtml) was searched for the flanking sequence positions. Of the 340 FSTs analyzed in total, 239 (70.3%) of the insertions were inside or within 600 bp of annotated genes (see Table S1). Of these 239 insertions, 167 (69.9%) of them were in known classes of genes (some examples are presented in Table 5) and 72 (30.1%) of them were in hypothetical genes. Of the 239 insertions into genic regions, 165 (69%) were within the coding region. The fraction of insertion into genes is expected to grow as gene annotation becomes more complete. Considering that annotated genes constitute less than 30% of the genome, these data indicate a preference of the *dSpm* element to insert into genes. A detailed analysis was carried out for 132 *dSpm* insertions into annotated genes for which predicted gene structures are available in the TIGR database, but there were no obvious preferences for *dSpm* element insertion into exons, introns, 5′-UTRs or 3′-UTRs (data not shown). However, it is possible that there may be weak preferences which could be detected using a larger dataset of FSTs.

Transposant siblings of the same T_3 family carry independent insertions

If transposition occurs early in development, then all the selected stable transposant siblings will carry the same insertion event, and hence each T_2 parent will be the source of a single independent transposant. If transposition occurs late in development, several independent transposants might be recovered from one T_2 parent. The flanking sequences from transposant siblings from the same T_3 family revealed independent insertions in most of the transposants analyzed. Sequences were analyzed by BLAST (National Center for Biotechnology Information, Bethesda, MD, USA) searches for two to four sibling transposants originating from the same T_3 family. We found that siblings from 86% of T_3 families (93 of 108) represented at least two independent transposition events (Table 6). Of these 93 T_3 families, 76

Table 6 T_3 families with independent defective suppressor mutator (*dSpm*) insertions in siblings

Transformant ID	T_3 families with more than one sibling	T_3 families with independent insertion in siblings	% siblings with different events
1	21	18	85.7
2	6	5	83.3
7	5	3	60
8	6	5	83.3
12	8	8	100
20	7	6	85.7
21	21	21	100
25	34	27	79.4
Total	108	93	86.11

(70%) had all siblings with different insertions (two of two siblings with different insertions for 56 families, three of three for 14 families, and four of four for six families), 13 had two of three siblings with different insertions, and four had two of four siblings with different insertions. These results suggest that *dSpm* transposition in this system occurs relatively late in development and that multiple independent insertion lines can be recovered from a single T_2 heterozygous parent.

Discussion

Development of an efficient En/Spm*-based gene tagging system in rice*

We have efficiently developed a novel transposon tagging system for rice based on the *Spm-dSpm* family of transposable elements. The maize *Spm-dSpm* family of transposable element was selected because studies in Arabidopsis have reported high frequencies of excision, integration and independent transposition events, which make it suitable for large-scale tagging in a heterologous system (Aarts *et al.*, 1995; Tissier *et al.*, 1999). An earlier study indicated that the use of the *En/Spm* system in rice may not be feasible as low frequencies of transposition were reported (Greco *et al.*, 2004). It was suggested that the low transposition frequency could be attributable to the influence of endogenous, autonomous, *En/Spm*-related elements present in the rice genome or attributable to the limiting amount of TnpA. In maize, the TnpA transcript is 100-fold more abundant than the TnpD transcript (Pereira *et al.*, 1986), whereas in rice the TnpD transcript appears to be more abundant (Greco *et al.*, 2004). In our study, we also observed a lower abundance of TnpA transcript when compared with TnpD. Although we observed lower levels of TnpA transcript in our rice lines, the *dSpm* element used in our study transposed efficiently. This indicates that the lower levels of transposase cannot be a major factor affecting *dSpm* transposition in rice. Instead, we believe that the low frequency of transposition in the study of Greco *et al.* (2004) might have been a result of the short termini of the *I/dSpm* element used (267 bp at the 5′ end and 640 bp at the 3′ end). Because our system utilizes a *dSpm* element with much longer terminal sequences (1014 bp at the 5′ end and 1193 bp at the 3′ end), the differences between our results and those of Greco *et al.* can be explained on the basis of the differences in the termini of the transposons used in the two studies. In maize, a naturally occurring 2.2-kb *dSpm* element from bronze-mutable13 (*bz-m13*) has been shown to have higher germinal transposition frequencies when compared with its smaller size deletion derivatives (Schiefelbein *et al.*, 1985). The size of the *dSpm* element used in our study is similar to that of the naturally occurring 2.2-kb *dSpm* element.

Plant transposable elements of the hAT *(hobo*, *Ac* and *Tam3*) and CACTA (conserved 5′-CACTA-3′ ends) superfamily, including *Ac-Ds* and *En/Spm* are known to transpose preferentially to linked sites. Hence genome-wide coverage can be achieved by introducing transposons carrying T-DNAs throughout the genome (Bancroft and Dean, 1993; Fedoroff and Smith, 1993; Ito *et al.*, 1999; James *et al.*, 1995; Muskett *et al.*, 2003; Nishal *et al.*, 2005). This approach requires hundreds of mapped transposon-carrying T-DNA lines to ensure that the entire genome is covered. An alternate approach for genome-wide coverage uses a small number of starter lines but enriches for unlinked transposition events by selecting against the donor sites (Parinov *et al.*, 1999; Sundaresan *et al.*, 1995; Tissier *et al.*, 1999).

We have previously demonstrated that the *GFP* gene is an efficient counter-selection marker to enrich for unlinked transposition (Kolesnik *et al.*, 2004) and has advantages over other markers used such as GUS (Nakagawa *et al.*, 2000) or *SU1* (Chin *et al.*, 1999). Another major advantage of using *GFP* in this system is that it can be used to distinguish the heterozygous T_2 seedlings based on the intensity of GFP fluorescence (Table 3), thereby reducing labor and space requirements by eliminating homozygous seedlings at early stages of the mutagenesis.

Currently, most heterologous transposon tagging systems in rice use the Basta Resistance (*BAR*) gene as the positive selection marker to select for the presence of the transposon. Unfortunately, BAR selection often produces escapes that can be as high as 20% (Kolesnik *et al.*, 2004). The use of the *DsRed* marker in this study has many advantages over *BAR*. First, only seedlings that carry the *DsRed* gene show red fluorescence at the excitation wavelength. Our results showed that all the selected DsRed plants based on red fluorescence contained the *DsRed* gene and the *dSpm* element, indicating a lack of escapes in the DsRed selection system. Secondly, unlike the Basta® screen (AgrEvo, Marysville, CA, USA), which involves planting of the GFP^- seedlings in soil before spraying with Basta, the screen for DsRed can be carried

out simultaneously along with the screen for GFP. Thirdly, transgenic plants carrying DsRed do not require stringent containment procedures, unlike those carrying BAR, which may be a cause of concern regarding unintentional transmission of herbicide resistance. Lastly, the zygosity of insertion lines carrying the *dSpm-DsRed* insertion can be inferred based on the intensity of red fluorescence, which can be used to accelerate segregation analysis of mutants.

Comparison of the En/Spm *system and the* Ac-Ds *system in rice*

In rice, the *Ac-Ds* system for gene tagging has been reported by many groups (Chin *et al.*, 1999; Enoki *et al.*, 1999; Greco *et al.*, 2001, 2003; Kim *et al.*, 2004; Kohli *et al.*, 2001; Upadhyaya *et al.*, 2002). However, their strategies and the marker genes they used are quite different from those in our *En/Spm* system. Hence we have used the results from our previously reported *Ac-Ds* tagging system, which is closest to our current study with the *En/Spm* system, to compare the transposition patterns of these two systems. Although the number of FSTs and the launching pad position in the two systems are quite different, useful information can still be obtained about the behavior of these two different transposon families in rice. We reported that, with the *Ac-Ds* system in rice, the *Ds* preferentially inserted into genic regions, representing about 72% of *Ds* insertions analyzed (144 of 200) (Kolesnik *et al.*, 2004). A comparable preference was observed for *dSpm* in this study, representing 70% of *dSpm* insertions (239 of 340). However, a detailed examination of the insertion sites revealed some interesting differences between the two systems. The results from this study indicate that the *Spm-dSpm* system generates more unlinked transpositions than *Ac-Ds*, and therefore is likely to be more efficient for genome-wide coverage. This conclusion is supported by three observations. First, the intrachromosomal insertion rate observed for the *dSpm* element was much lower (12%) than the intrachromosomal insertion rate for *Ds* (24%) (Kolesnik *et al.*, 2004). Secondly, about 1.9% of the new *dSpm* insertions were inserted within 4 cM of the original donor locus, whereas, in the *Ac-Ds* system, 6% of the insertions were within 4 cM of the original donor locus (Kolesnik *et al.*, 2004). Thirdly, the frequency of *dSpm* insertions within the T-DNA was around 3.4%, which is a substantial improvement over the *Ac-Ds* system, where approximately 12% of insertions were found to be within the T-DNA (Kolesnik *et al.*, 2004).

Using the *Ac-Ds* system, we previously reported the occurrence of some 'hot spots' for *Ds* insertion in the rice genome (Kolesnik *et al.*, 2004). The most pronounced hot spot was in chromosome 7, and about 4% (42 of 979) of the *Ds* insertions analyzed in that study were in a 40-kb region in chromosome 7 (76.6 cM). None of the *dSpm* insertions analyzed here was within that region, suggesting that the same chromosomal region is not a hot spot for *dSpm* insertions. In our study, no hot spots could be detected for *dSpm* insertion in the rice genome; however, there may be a weak preference for *dSpm* insertion in a 10-kb region on chromosome 1 (73.4 cM), representing 0.87% or three of 341 insertions. These three insertions originated from two different transformants, numbers 22 and 25, which were on chromosomes 1 (32.4 cM) and 7 (103.4 cM), respectively. To summarize, our results suggest that *Ds* and *dSpm* exhibit different preferences for insertion in the rice genome, and hence different genome coverage is likely to be achieved using these elements.

Application for high-throughput insertion line production

The ease of selecting stable transformants using the fluorescence marker genes *GFP* and *DsRed* without antibiotics or herbicides makes our system suitable for high-throughput insertional mutagenesis. Stable *dSpm* insertions can be obtained from the progenies arising from the self-pollination of plants that are heterozygous for the T-DNA locus. For high-throughput insertion line production, fluorescent sorting of GFP^{-} and $DsRed^{+}$ seedlings could potentially be automated. Transposant screening using seedlings in 96-well plates and Kodak Image Station 2000 MM (Figure 6) can be adapted with other robotic components.

In our transformation experiments, we obtained approximately 50% of transformants with single T-DNA insertions. On average, a single T_1 transformant can produce about 400 seeds and can potentially yield 200 T_2 heterozygous families for transposant screening. Based on our pilot-scale experiments, a single transformant carrying single-locus T-DNA can generate about 80 unique insertion lines, assuming a 40% frequency of families carrying unlinked germinal events. Hence about 500 single-copy transformants can be used to generate up to approximately 40 000 independent insertion lines. This amplification factor is achieved simply through a generation of propagation by selfing, followed by fluorescence screening. These numbers could potentially be increased further by selecting two or more siblings from each T_3 family, as most T_3 siblings appeared to be independent transpositions. Because the approach relies on a limited number of primary transformants, it can be easily adapted to crop plants for which large-scale T-DNA mutagenesis is impractical, including other subspecies of rice such as *indica*. In conclusion, we have established a new insertional mutagenesis system in rice using the *En/Spm* transposon, which can usefully complement other mutagens such as *Tos17*, T-DNA and *Ac-Ds*, and fill gaps left by these elements in the rice genome.

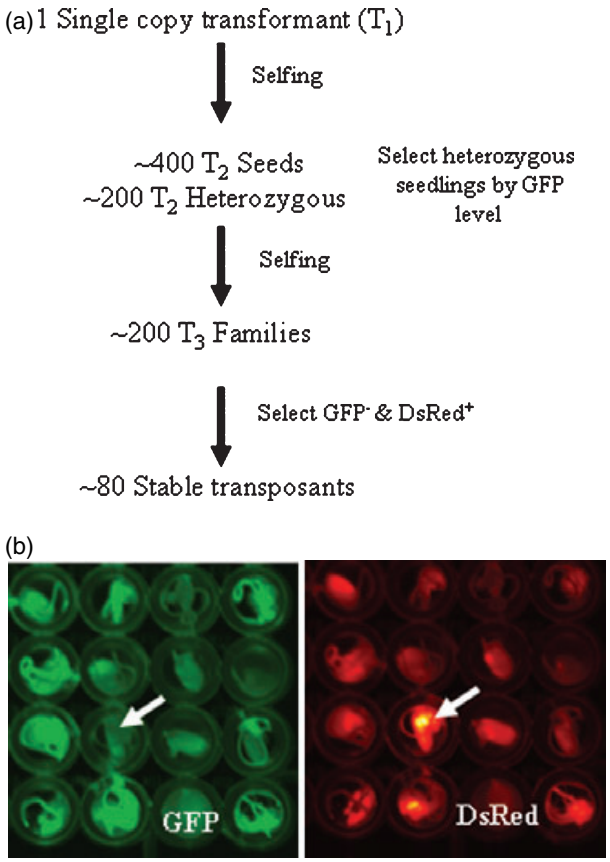

Figure 6. (a) Strategy for large-scale insertion line generation. (b) High-throughput transposant screening using seedlings in 96-well plates and Kodak Image Station 2000 MM. Arrows indicate a transposant that is negative for green fluorescent protein (GFP^-; left panel) and positive for *Discosoma sp.* Red Fluorescence Protein ($DsRed^+$; right panel).

Experimental procedures

Generation of the tagging vector

Binary vector. The T-DNA binary vector pCAMBIA1300 (kindly provided by R. Jefferson, CAMBIA, Canberra, Australia) was used as the base vector for constructing the tagging vector. This vector also carries a CaMV 35S promoter-driven hygromycin-resistant gene towards the left border for selecting transformants.

GFP *negative selection marker.* The expression cassette carrying Ubiquitin promoter-sGFP-Nopaline synthase terminator (UbiP-sGFP-NosT) was constructed by ligating a *Kpn*I-*Sac*I fragment containing sGFP-NosT (Chiu *et al.*, 1996) to pCAMBIA1300-UbiP to generate the binary vector pSK50.

dSpm-*DsRed.* A full-length suppressor-mutator (N. Fedoroff, Pennsylvania State University, PA, USA) *Xho*I fragment was treated with Klenow (New England Biolabs, Ipswich, MA, USA) and cloned into the Klenow-treated *Pst*I site of pMUCBS, a low-copy-number derivative of pBR322. Then the new defective *Spm* (*dSpm*) element was created by deleting the internal fragment bordered by the most proximal and most distal *Pst*I sites of the full-length *Spm* element followed by self-ligation. The unique *Pst*I site within the *dSpm* element was used to clone the expression cassette carrying the rice Ubi11 promoter driving DsRed, the positive selection marker, and Nos termination signal, to generate p*dSpm*-DsRed.

Immobilized Spm-*transposase cassette.* The full-length *Spm* element was digested with *BssH*II and treated with the Klenow fragment. The pMUCBS which carries the CaMV 35S promoter and Nos terminator was linearized using *Sma*I and the immobilized Spm was ligated to it to generate 35S-imm.Spm. The 35S-imm.Spm transposase fragment could then be excised as a *Cla*I fragment.

The final construct. The *Cla*I fragment carrying 35S-imm.Spm was then inserted into the *Cla*I site of p*dSpm*-DsRed to generate p*dSpm*-imm.Spm. This was *Not*I digested and Klenow-treated, ligated to pSK50 DNA that had been *Hind*III digested and filled-in using Klenow, to generate the tagging vector.

Agrobacterium-*mediated rice transformation*

Scutellum-derived embryogenic calli were generated from mature seeds of *O. sativa* ssp. *japonica* cv. Nipponbare. The calli were co-cultivated with *Agrobacterium* strain AGL1 (Lazo *et al.*, 1990) harboring the binary vector, as described by Hiei *et al.* (1994) with slight modifications. The transformed calli were selected on 50 mg L^{-1} hygromycin B and 200 mg L^{-1} timentin. The transgenic calli expressing *GFP* were regenerated into plantlets in basal Nutrient 6 macro elements and B5 micro elements (NB) medium. The primary transformants (T_1) were transferred to the glasshouse and seeds were harvested on maturity.

GFP *and* DsRed *expression analysis*

GFP and *DsRed* expression in transformed calli, regenerated plants, mature seeds or germinating seedlings were observed using a Zeiss SV11 fluorescence microscope (Zeiss, Hamburg, Germany) with appropriate filters for GFP and DsRed fluorescence.

Determination of the zygosity of progenies

The zygosity of seedlings was determined based on the intensity GFP fluorescence levels. Mature seeds were soaked in water for 4–6 days and the germinating seedlings were examined for GFP fluorescence levels using the Zeiss SV11 fluorescence microscope. Then the selected heterozygous T_2 seedlings were planted in the glasshouse for transposant generation.

Screening for stable transposants

The mature seeds were soaked in water for 4–6 days and then the emerging seedlings were screened for GFP fluorescence, using the Green Fluorescent Protein Macro Detector Set (Model GFP-MDS-20/BB; Biological Laboratory Equipment Maintenance and Service Ltd, Budapest, Hungary). The selected GFP-negative seedlings were then checked for DsRed fluorescence using the Zeiss SV11 fluorescence microscope. The lines with insertions into the *GFP* gene were eliminated either by looking for *GFP* sectors or by means of PCR screen to select out insertions into the *GFP* gene.

Molecular analysis of plants

Leaf samples were collected from the transformants or transposants after transfer into the glasshouse. The genomic DNA was isolated using the Qiagen DNeasy kit (Qiagen, Cologne, Germany). All the regenerated transformants were checked by PCR with hygromycin phosphotransferase (HPH)-specific primers (HPH-F: 5′-GCCCATCTTAGCCAGACG-3′ and HPH-R: 5′-CAACCAAGCTCTGATAG-3′), DsRed-specific primers (DsRed-F: 5′-GCTGAAGGTGACCAAGGG-3′ and DsRed-R: 5′-GATGTCCAGCTTGGAGTC-3′) and *Spm*-transposase-specific primers (Spm-F: 5′-CTAGAAGCGAGTCGTCCAC-3′ and Spm-R: 5′-CGACCTTCGTGATCTGGATC-3′). All PCR reactions were performed with approximately 50–100 ng of genomic DNA under standard amplification conditions.

Excision PCR

The excision of the *dSpm* element from the donor locus was detected using the primers designed on either side of the *dSpm* element in the T-DNA construct (Ex-5: 5′-TAAGTTGGGTAACGCCAGGGT-3′ and Ex-3: 5′-CATCTATGTTACTAGATCGGG-3′). This combination of primers will produce a band only if there is an excision amplifying the EDS. The untransposed *dSpm* elements were detected by the amplification of the full donor site (FDS), by using the combination of primers at the 5′ and 3′ ends of the *dSpm* element.

RNA isolation and Tnp transcript expression analysis

Total RNA was isolated from 10-day-old seedlings using the Qiagen RNeasy kit. Semiquantitative RT-PCR was performed using cDNA synthesized from total RNA with Superscript reverse transcriptase (Invitrogen, Grand Island, New York, NY, USA). The *En/Spm* specific primers were designed based on Masson *et al.* (1989) to anneal at positions in the *En/Spm* sequences. We used the same combinations of primers used by Greco *et al.* (2004) in rice: primer 1 (560-Forward), primer 7 (3781-Forward), primer 11 (5299-Reverse), primer 12 (6059-Forward), primer 13 (6402-Reverse) and primer 14 (7761-Reverse). Aliquots of amplified products were separated on agarose gel stained with Gel Star (Cambrex Bio Science Rockland, Inc., Rockland, ME, USA).

Genomic DNA isolation and Southern blot hybridization

Genomic DNA from leaf tissue was isolated as described by Dellaporta *et al.* (1983). For each sample, 4–5 μg of genomic DNA was digested with the appropriate enzymes, fractionated on 0.8% agarose gel and transferred onto a Hybond-N^+ membrane (Amersham, Sunnyvale, CA, USA). Blots were hybridized with ^{32}PdCTP-labeled probes in Rapid Hyb solution (Amersham) at 65°C. After hybridization, membranes were washed twice with 2 × saline sodium citrate (SSC) and 0.1% sodium dodecyl sulfate (SDS) for 15 min, then twice with 0.5 × SSC and 0.1% SDS for 15 min at 65°C, and finally twice with 0.1 × SSC and 0.1% SDS at 65°C. The ^{32}PdCTP-labeled probes were prepared using the Redi Prime II DNA labeling kit (Amersham) according to the manufacturer's protocols.

Analysis of T-DNA and dSpm *flanking sequences*

Fragments flanking the T-DNA inserts in a parental line and the *dSpm* element in a transposant line were amplified by TAIL-PCR as described by Liu and Whittier (1995). The following nested specific primers were designed for the T-DNA right border and *dSpm* 5′ and 3′ ends were used:

dSpm 5′
Spm5-1: 5′-GGTGTGGAAAAACCCACACTC-3′
Spm5-2: 5′-GCGTCGGTTTCATCGGGACC-3′
Spm5-3: 5′-CTCTTTAATTAACTGACACTCC-3′
dSpm 3′
Spm3-1: 5′-GTCGGTCCCCACACTTCTATACG-3′
Spm3-2: 5′-GAGCGTCCATTTTAGAGTGAC-3′
Spm3-3: 5′-GTGTGGGGTTTTGGCCGACAC-3′
T-DNA right border
RB-1: 5′-GATCGCCCTTCCCAACAGT-3′
RB-2: 5′-CTGAATGGCGAATGCTAGAGC-3′
RB-3: 5′-GTTTGACAGGATATATTGGCGG-3′
Successfully used arbitrary primers were:
AD-11: 5′-NCASGAWAGNCSWCAA-3′
AD-13: 5′-NTSGASNTCNGAATCA-3′

Purification and sequencing of flanking DNA

The TAIL-PCR products were resolved on 1.75% agarose gels and extracted from gels using the QiaQuick Gel Extraction Kit and QiaQuick 96 PCR purification kit (Qiagen). DNA was sequenced using the ABI Prism Big-Dye Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems, Foster City, CA, USA).

Flanking sequences were sent to National Center for Biotechnology Information (NCBI) GenBank and the TIGR database to search for similarities. Two algorithms, BLASTN [cut-off E-value-100 to obtain identical hits on rice Bacterial Artificial Chromosomes (BACs) and P1-derived Artificial Chromosomes (PACs); NCBI, Bethesda, MD, USA] and BLASTX (cut-off E-value-10; NCBI), were used to analyze each sequence (Altschul *et al.*, 1990). To map *Ds* insertions, a BAC/PAC *in silico* integrated map was utilized from the TIGR/RGP database.

Acknowledgements

We are grateful to Anna Persson for helping with DNA extraction. We also thank Asuman Buzkan, Patrick Hogan and Cameron Johnson for their valuable help, and Farm Saeteurn, Eva Yang and Jodie Mae Bernardino for their assistance with planting and harvesting rice. We are grateful to Dr Nina Fedoroff for the *Spm* clone. This research was supported by National Science Foundation Plant Genome Research Program award # 0211924.

Supplementary Material

The following supplementary material is available for this article online:

Table S1 List of defective suppressor mutator (*dSpm*) insertions into annotated rice genes. The *dSpm* insertion lines and their flanking sequences described in this paper are available publicly through our laboratory web site http://sundarlab.ucdavis.edu/rice/blast/blast.html

This material is available as part of the online article from http://www.blackwell-synergy.com

References

Aarts, M.G., Corzaan, P., Stiekema, W.J. and Pereira, A. (1995) A two element *Enhancer-Inhibitor* transposon system in *Arabidopsis thaliana*. *Mol. Gen. Genet.* **247**, 555–564.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic Local Alignment Search Tool. *J. Mol. Biol.* **215**, 403–410.

Bancroft, I. and Dean, C. (1993) Transposition pattern of the maize element *Ds* in *Arabidopsis thaliana*. *Genetics*, **134**, 1221–1229.

Bennetzen, J.L. and Ma, J. (2003) The genetic colinearity of rice and other cereals on the basis of genomic sequence analysis. *Curr. Opin. Plant Biol.* **6**, 128–133.

Chin, H.G., Choe, M.S., Lee, S.H. *et al.* (1999) Molecular analysis of rice plants harboring an *Ac/Ds* transposable elements mediated gene trapping system. *Plant J.* **19**, 615–623.

Chiu, W., Niwa, Y., Zeng, W., Hirano, T., Kobayashi, H. and Sheen, J. (1996) Engineered GFP as a vital reporter in plants. *Curr. Biol.* **6**, 325–330.

Christensen, A.H. and Quail, P.H. (1996) Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res.* **5**, 213–218.

Dellaporta, S.L., Wood, J. and Hicks, J.B. (1983) A plant DNA minipreparation: version 2. *Plant Mol. Biol. Rep.* **1**, 19–22.

Delseny, M. (2004) Re-evaluating the relevance of ancestral shared synteny as a tool for crop improvement. *Curr. Opin. Plant Biol.* **7**, 126–131.

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. and Smith, D.L. (1993) A versatile system for detecting transposition in *Arabidopsis*. *Plant J.* **3**, 273–289.

Gale, M.D. and Devos, K.M. (1998) Comparative genetics in the grasses. *Proc. Natl Acad. Sci. USA*, **95**, 1971–1974.

Goff, S.A., Ricke, D., Tien-Hung, L., *et al.* (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). *Science*, **296**, 92–100.

Greco, R., Ouwerkerk, P.B., Taal, A.J., Favalli, C., Beguiristain, T., Puigdomenech, P., Colombo, L., Hoge, J.H. and Pereira, A. (2001) Early and multiple *Ac* transpositions in rice suitable for efficient insertional mutagenesis. *Plant Mol. Biol.* **46**, 215–227.

Greco, R., Ouwerkerk, P.B.F., de Kam, R.J., Salloud, C., Favalli, C., Colombo, L., Guiderdoni, E., Meijer, A.H., Hoge, J.H.C. and Pereira, A. (2003) Transpositional behaviour of an *Ac/Ds* system for reverse genetics in rice. *Theor. Appl. Genet.* **108**, 10–24.

Greco, R., Ouwerkerk, P.B., Taal, A.J., Sallaud, C., Guiderdoni, E., Meijer, A.H., Hoge, J.H. and Pereira, A. (2004) Transcription and somatic transposition of the maize *En/Spm* transposon system in rice. *Mol. Gen. Genet.* **270**, 514–523.

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., Sugimoto, K., Otsuki, Y., Tsugawa, H. and Kanda, M. (1996) Retrotransposaons of rice involved in mutations induced by tissue culture. *Proc. Natl Acad. Sci. USA*, **93**, 7783–7788.

Ito, T., Seki, M., Hayashida, N., Shibata, D. and Shinozaki, K. (1999) Regional insertional mutagenesis of genes on Arabidopsis thaliana chromosome V using the *Ac/Ds* transposon in combination with a cDNA scanning method. *Plant J.* **17**, 433–444.

Ito, Y., Eiguchi, M. and Kurata, N. (2004) Establishment of an enhancer trap system with *Ds* and GUS for functional genomics in rice. *Mol. Genet. Genomics*, **271**, 639–650.

Izawa, T., Ohnishi, T., Nakano, T. *et al.* (1997) Transposon tagging in rice. *Plant Mol. Biol.* **35**, 219–229.

James, D.W., Jr, Lim, E., Keller, J., Plooy, I., Ralston, E. and Dooner, H.K. (1995) Directed tagging of the Arabidopsis *FATTY ACID ELONGATION1* (*FAE1*) gene with the maize transposon *Activator*. *Plant Cell*, **7**, 309–319.

Jeong, D.H., An, S., Kang, H.G., Moon, S., Han, 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.

Khush, G.S. (1997) Origin, dispersal, cultivation and variation of rice. *Plant Mol. Biol.* **35**, 25–34.

Kim, C.M., Piao, H.L., Park, S.J., *et al.* (2004) Rapid, large-scale generation of *Ds* transposant lines and analysis of the *Ds* insertion sites in rice. *Plant J.* **39**, 252–263.

Kohli, A., Xiong, J., Greco, R., Christou, P. and Pereira, A. (2001) Tagged Transcriptome Display (TTD) in indica rice using *Ac* transposition. *Mol. Genet. Genomics*, **266**, 1–11.

Kolesnik, T., Szeverenyi, I., Bachman, D., Kumar, C.S., Jiang, S., Ramamoorthy, R., Cai, M., Ma, Z.H., 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.

Lazo, G.R., Stein, P.A. and Ludwig, R.A. (1990) A DNA transformation-competent Arabidopsis genomic library in *Agrobacterium*. *Biotechnology*, **9**, 963–967.

Liu, Y.G. and Whittier, R.F. (1995) Thermal asymmetric interlaced PCR, automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics*, **25**, 674–681.

Marsch-Martinez, N., Greco, R., van Arkel, G., Herrera-Estrella, L. and Pereira, A. (2002) Activation tagging using the *En-I* maize transposon system in Arabidopsis. *Plant Physiol.* **129**, 1544–1556.

Masson, P. and Fedoroff, N.V. (1989) Mobility of the maize suppressor-*mutator* element in transgenic tobacco cells. *Proc. Natl Acad. Sci. USA*, **86**, 2219–2223.

Masson, P., Rutherford, G., Banks, J. A. and Federoff, N.V. (1989) Essential large transcripts of the maize *Spm* transposable element are generated by alternative splicing. *Cell*, **58**, 755–765.

Masson, P., Surosky, R., Kingsbury, J.A. and Fedoroff, N.V. (1987) Genetic and molecular analysis of the Spm-dependent *a-m2* allele of the maize *a* locus. *Genetics*, **117**, 117–137.

McClintock, B. (1954) Mutations in maize and chromosomal aberrations in meurospora. *Carnegie Inst. Wash. Year Book*, **56**, 254–260.

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.

Muskett, P.R., Clissold, L., Marocco, A., Springer, P.S., Martienssen, R. and Dean, C. (2003) A resource of mapped dissociation launch pads for targeted insertional mutagenesis in the Arabidopsis genome. *Plant Physiol.* **132**, 506–516.

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.

Nishal, B., Tantikanjana, T. and Sundaresan, V. (2005) An inducible targeted tagging system for localized saturation mutagenesis in Arabidopsis. *Plant Physiol.* **137**, 3–12.

Parinov, S., Sevugan, M., Ye, 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.

Pereira, A., Cuypers, H., Gierl, A., Schwarz-Sommers, Z. and Saedler, H. (1986) Molecular analysis of the *En/Spm* transposable element system of Zea mays. *EMBO J.* **5**, 835–841.

Peterson, P. (1953) A mutable *pale green* locus in maize. *Genetics*, **45**, 113–115.

Sallaud, C., Gay, C., Larmande, P. *et al.* (2004) High throughput T-DNA insertion mutagenesis in rice, a first step towards *in silico* reverse genetics. *Plant J.* **39**, 450–464.

Schiefelbein, J.W., Raboy, V., Fedoroff, N.V. and Nelson, O.E., Jr (1985) Deletions within a *defective suppressor-mutator* element in maize affect the frequency and developmental timing of its excision from the bronze locus. *Proc. Natl Acad. Sci. USA*, **82**, 4783–4787.

Speulman, E., Metz, P.L.J., van Arkel, G., 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.

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

Tissier, A.F., Marillonnet, S., Kimyuk, 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.

Upadhyaya, N.M., Zhou, X.R., Zhu, Q.H. *et al.* (2002) An *iAc/Ds* gene and enhancer trapping system for insertional mutagenesis in rice. *Funct. Plant Biol.* **29**, 547–559.

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

Wu, C., Li, X., Yuan, W *et al.* (2003) Development of enhancer trap lines for functional analysis of the rice genome. *Plant J.* **35**, 418–427.

Yu, J., Hu, S., Wang, J., *et al.* (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science*, **296**, 79–92.