Mol Gen Genomics (2006) 276:533–543
DOI 10.1007/s00438-006-0158-0

ORIGINAL PAPER

Use of the transposable element *Ac/Ds* in conjunction with *Spm/dSpm* for gene tagging allows extensive genome coverage with a limited number of starter lines: functional analysis of a four-element system in *Arabidopsis thaliana*

Priya Panjabi · Pradeep Kumar Burma · Deepak Pental

Received: 8 May 2006 / Accepted: 7 August 2006 / Published online: 27 September 2006

Abstract We have developed a novel four-element based gene tagging system in *Arabidopsis* to minimize the number of starter lines required to generate genome-wide insertions for saturation mutagenesis. In this system, the non-autonomous cassette, *Ds*(*dSpm*), comprises of both *Ds* and *dSpm* elements cloned one within the other along with appropriate selection markers to allow efficient monitoring of excision and re-integration of the transposons. *Trans*-activation of the outer borders (*Ds*) and selection against the negative selection marker (*iaaH*) linked to the cassette ensures unlinked spread of the *Ds*(*dSpm*) cassette from the initial site of integration of the T-DNA. This creates several launch pads within the genome from where the internal element (*dSpm*) can be subsequently mobilized to generate secondary insertions. In this study, starting from a single T-DNA integration we could spread the *Ds*(*dSpm*) cassette to 11 different locations over all the five chromosomes of *Arabidopsis*. The frequency of unlinked *Ds* transpositions in the F2 generation varied between 0.05 and 3.35%. Three of these lines were then deployed to *trans*-activate the internal *dSpm* element which led to the selection of 29 *dSpm* insertions. The study conclusively shows the feasibility of deploying *Ds* and the *dSpm* elements in a single construct for insertional mutagenesis.

Communicated by M.-A. Grandbastien.

P. Panjabi · P. K. Burma · D. Pental (✉)
Department of Genetics, University of Delhi South Campus, Benito Juarez Road, New Delhi 110021, India
e-mail: dpental@gmail.com

Keywords *Ac/Ds* · Four-element system · Insertional mutagenesis · *Spm/dSpm* · Transposons

Introduction

Assignment of function to the ever-expanding repertoire of sequence information is a major consideration in biological sciences. Another important area (in terms of functional genomics) is the identification and isolation of novel alleles distributed in the wild relatives and even distant species of crop plants, which could contribute towards cultivar improvement. Whether the aim is global assessment of gene function (as is being done in the model plant *Arabidopsis thaliana* and more recently in rice) or selective assessment of alleles for agronomic use, development of tools for functional analysis have critical importance for the success of such programmes.

Currently, several methods for deciphering gene function are deployed in plants. These include insertional mutagenesis, gene replacement, RNAi-based gene silencing (Chen et al. 2003; Guo et al. 2003), physical mutagenesis using fast neutrons (Li et al. 2001) and chemical mutagenesis or TILLING (McCallum et al. 2000; Till et al. 2003; Slade et al. 2005). Of these, insertional mutagenesis is one of the most favored techniques as it requires no prior sequence information of the gene and can be suitably modified to generate dominant gain-in-function mutations by activation tagging (Wilson et al. 1996; Marsch-Martinez et al. 2002) or engineered to function as a gene trap system (Sundaresan et al. 1995; Smith and Fedoroff 1995; Kiegle et al. 2000).

Two-element based transposon systems utilizing maize transposable elements *Ac/Ds* and *Spm/dSpm*

have been suitably tailored for efficient gene isolation and disruption in plants (Hehl and Baker 1989; Masson and Fedoroff 1989; Masterson et al. 1989; Altmann et al. 1992; Bancroft et al. 1992). As both the transposable elements show a preference towards linked transpositions, negative selection markers were incorporated to screen for unlinked transposition events (Sundaresan et al. 1995; Tissier et al. 1999). This allowed for maximum coverage of the genome using fewer starter lines. In all the transposon-based systems described till date, only the first transposition event can be phenotypically monitored. A system that could help in monitoring subsequent jumps phenotypically after the initial transposition event would be a major advancement in gene tagging. This would make saturation mutagenesis more efficient as well as allow tagging genes or alleles of interest by using a combination of unlinked and linked transpositions. This system would be particularly helpful in plant species where generating a large collection of the initial T-DNA lines is difficult due to poor transformation frequencies.

We had earlier proposed a four-element based transposon tagging system which would allow both linked and unlinked jumps in two successive steps (Phogat et al. 2000). The four-element system uses a non-autonomous cassette, designated *Ds*(*dSpm*) or *dSpm*(*Ds*), which harbors both the *Ds* and *dSpm* border elements cloned one within the other along with appropriate marker genes to monitor the excision and re-integration of the two transposable elements. The system utilizes a negative selection marker and the outer non-autonomous element for the unlinked scattering of the transposon construct from the site of the initial T-DNA integration. This would lead to the generation of several launch pads at different locations in the genome from where the internal element can subsequently be mobilized for either generating localized mutagenesis (for allele-specific tagging) or for further scattering of the cassette for saturation mutagenesis. As the non-autonomous cassette would harbor both the *Ds* and the *dSpm* elements, their *trans*-activation would require their individual transposase source (*Ac* transposase for *Ds* and *Spm* transposase for *dSpm* mobilization). The system has been therefore termed as a 'Four-element' system. The essence of the system lies in the fact that it enables phenotypic selection of both the unlinked transpositions of the outer borders and the subsequent jumps of the internal element (*dSpm*) due to the presence of appropriate marker genes. In this study, we report the development of a *Ds*(*dSpm*) construct and its functionality in *A. thaliana* to show that the two transposable elements can be used in conjunction with gene tagging.

Materials and methods

Development of the pPZP100: *Ds*(*dSpm*) construct

The *Ds* cassette (~1.8 kbp *Ds*5′ and ~240 bp *Ds*3′) containing *nptII* gene (under the 1′ promoter) was removed from the vector pWS32 (Genetrap *Ds*, described by Sundaresan et al. 1995) and transferred to the low copy vector pMuCBs (received from V. Sundaresan). The *dSpm* borders (1.487 kbp *dSpm*5′ and 740 bp *dSpm*3′) generated by *Nsi*I digestion of the full-length *Spm* element were inserted between the 35S promoter and the coding region of the *aadA* gene in the vector, pLitmus38 (NEB). The *aadA*(*dSpm*) cassette was then placed between the *Ds* borders, alongside the *nptII* gene. Initially, problems were encountered in the introduction of the *aadA*(*dSpm*) fragment within the *Ds* borders. To circumvent this, the *Ds*3′ border was removed from the *Ds*5′-*nptII*-*Ds*3′ cassette following which the *aadA*(*dSpm*) fragment could be easily cloned. The *Ds*3′ border was then introduced as a separate fragment to generate the cassette *Ds*5′-*nptII*-*aadA*(*dSpm*)-*Ds*3′ (plasmid pMWSSD-*Ds*3′). The entire assembly was done such that the *Ds* cassette could be retrieved as a *Kpn*I fragment from the plasmid pMuCBs.

The 2′-*iaaH*-pA fragment was removed from plasmid pAJ6 (received from V. Sundaresan) and inserted into the *Kpn*I-*Sac*I site in the MCS of binary vector pPZP100 (Hajdukiewicz et al. 1994). Subsequently, the *Kpn*I fragment from plasmid pMWSSD-*Ds*3′ harboring the *Ds* cassette was introduced to generate the binary vector pPZP100: *Ds*(*dSpm*). The region within the T-DNA borders of the binary vector pPZP100: *Ds*(*dSpm*) is schematically represented in Fig. 1.

Development of *Ds*(*dSpm*) transgenics

In planta transformation (Bechtold et al. 1993) of *Arabidopsis thaliana* ecotype Columbia was carried out using *Agrobacterium tumefaciens* strain GV3101 harboring the binary vector pPZP100: *Ds*(*dSpm*). Transformed lines were identified by germination on medium (1× MS salts, 1× B5 vitamins, 1% sucrose and 0.8% Difco agar) supplemented with 50 mg/l kanamycin and 200 mg/l of the bacteriostatic agent cefotaxime. The plates were placed at 4°C for 3 days after which they were shifted to the growth room (22°C with 10 h light/14 h night photoperiod). The Kan^R seedlings were transferred to fresh selection plates and subsequently into individual pots. Homozygous transgenic *Ds*(*dSpm*) lines were developed by selfing and used subsequently for crosses with transposase containing transgenic lines.

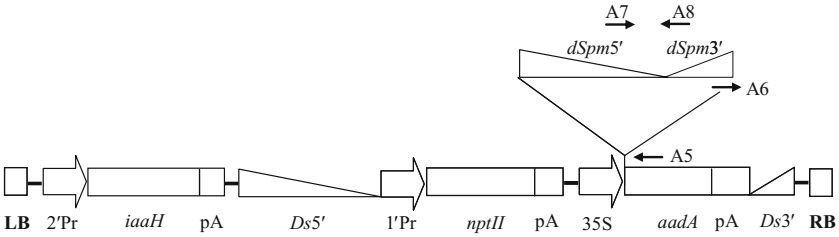

Fig. 1 Details of the *Ds*(*dSpm*) construct in binary vector pPZP100. The sequences cloned between the binary vector T-DNA borders have been shown. The *Ds* element contains an *aadA* gene (conferring resistance to streptomycin and spectinomycin) and an *nptII* gene (conferring resistance to kanamycin). Kanamycin resistance serves as a selectable marker for transgenic plants harboring the *Ds*(*dSpm*) construct and also as a re-integration marker for the *Ds* cassette following unlinked transposition while the *aadA* gene serves as an excision marker for *dSpm*. The *iaaH* gene codes for indole acetic acid hydrolase and acts as a negative selection marker to select for unlinked *Ds* transpositions. The *arrows* depict the position of the different primers used to characterize the *dSpm* excision events. Primers A5 and A6 are specific to pre-excision (~460 bp), while A7 and A8 are specific to *dSpm* borders (~750 bp)

Transposase containing lines used for *in-trans* activation of *Ds*(*dSpm*) construct

Two independent, homozygous *Ac* transposase lines were used in the current study: sAc line harboring the s*ΔNae*-IAc element (Bancroft et al. 1992) and Ac2 line harboring 35 S::Ac transposase fusion lines (Sundaresan et al. 1995). Both the lines have the *Ac*TPase cassette linked to *nptII* and *iaaH* genes, but while in Ac2 the *Ac* TPase expression is driven by the 35 S promoter, in sAc, *Ac* is expressed from its own promoter (a 537 bp deletion has been created within the 5′-UTL of *Ac*). The Ac2 line has been shown to give >30% germinal excision frequency but a low re-insertion frequency of *Ds* (Swinburne et al. 1992; Long et al. 1993). The sAc line on the other hand has been reported to give a lower frequency of germinal excision (~5%; Bancroft et al. 1992), but a higher percent of re-insertion events (~50%; Bancroft and Dean 1993).

For the *Spm* transposase, *A. thaliana* (Col) transgenic line (Spm404) harboring the construct 35 S*TnpA*-35 S*TnpD-bar-iaaH* (received from Monika Frey) was used.

Mapping of T-DNA and transposon integration sites

Genomic sequences flanking the insertion lines were amplified using Directional Genome Walking PCR (Mishra et al. 2002) or TAIL-PCR (Liu et al. 1995) with slight modifications in either of the protocols. Genomic DNA was isolated by CTAB method (Rogers and Bendich 1994) or by using DNeasy Plant Mini kit (Qiagen). PCR reactions were carried out using arbitrary primers and gene-specific nested primers. The arbitrary primers AD1, AD2, AD5, AD6 and AD8 (Liu et al. 1995; Liu and Whittier 1995; Tsugeki et al. 1996) and RWP1-RWP4 (Mishra et al. 2002) were deployed for TAIL-PCR and Genome Walking PCR, respectively. Each gene-specific primer set comprised of three nested primers, one each to be used in the three successive PCR reactions. The primary-specific primer of each set was biotinylated so that the same primer sets could be deployed for Genome walking PCR. Two sets of Ds3′ and Ds5′ end-specific primers (previously described by Parinov et al. 1999) were used. The *iaaH* specific primers consisted of iaaH1 (5′-CCgCAA ACC ATC CCA gTC TgT A-3′), iaaH 1.2 (5′-gCT TCA CAA CgC gCT ATC AgA g-3′) and iaaH1.3 (5′-TCC gTT TCA ggT gTT CTA ggC T-3′). The LB set consisted of RRB1 (5′-ATT CAA TTC ggC gTT AAT TCA gTA CA-3′), RRB1.2 (5′-ACg TCC gCA ATg TgT TAT TAA gTT g-3′) and RRB1.3 (5′-AAg CgT CAA TTT gTT TAC ACC ACA A-3′). The *dSpm3′*-specific primers included the following: dSpm3-1a (5′-ggC TTA TTT CAg TAA gAg Tg-3′), dSpm3-2a (5′-gTA AgA gTg Tgg ggT TTT ggC C-3′) and dSpm3-3a (5′-Tgg CCg ACA CTC CTT ACC TTT T-3′). The *dSpm* 5′-specific set comprised of dSpm5-1 (5′-TAT ACT ggA Cgg Cgg gCA gg-3′), dSpm5-2 (5′-ggA gAg AAg AAg CAC gAC ggC-3′) and dSpm5-3 (5′-Agg ATT ggg gAA TTT Agg gTA ACA T-3′).

The amplified PCR products obtained with either method were electrophoresed on agarose gels; the band of interest was eluted (GFX™ PCR, DNA and Gel band purification kit, Amersham Biosciences) and sequenced using ABI Prism™ 310 Genetic Analyzer (Applied Biosystems). The flanking sequences obtained were then subjected to BLASTN search of the NCBI Genbank database to map the insertions on the *Arabidopsis* genome.

Molecular characterization of the *dSpm* excision events

PCR amplifications were carried out on tissue samples as described by Klimyuk et al. 1993. The *dSpm* pre-excision specific primers, A5 (5′-CgC TCg ATg ACg CCA ACT

ACC TC-3′) and A6 (5′- gCC CAg gTA gCT TAC TgA TgT-3′), were designed to amplify approximately 460 bp product (Fig. 1). Primers used for analyzing re-integration of the *dSpm*, A7 (5′-CAA CgA ATC CTg CAT ACC CAA A-3′) and A8 (5′-ACC Tgg gCg CAT ATA AgA gTg Tgg-3′), were designed from within the *dSpm* borders to amplify approximately 750 bp product (Fig. 1).

The PCR reaction mix comprised of the specific primers (50 pmoles each), 200 µM dNTPs (NEB), 1× *Taq* polymerase buffer (Finnzyme), 1 U of Finnzyme *Taq* polymerase and 0.05% nonidet in a final volume of 50 µl. The PCR parameters were as follows: initial denaturation at 94°C for 5 min followed by 30 cycles of 94°C for 1 min, 55°C for 1 min, 72°C for 1 min. A final extension at 72°C was given for 5 min.

Results

Experimental strategy

The four-element system is supposed to function in two-steps (Fig. 2). In the first step, *Ds*(*dSpm*) and *Ac* transposase lines were crossed to *trans*-activate the *Ds* cassette. Seeds from selfed F1 lines were screened for the presence of unlinked *Ds* transpositions based on Kan^R-NAM^R phenotype. Since both *Ac* and the *Ds*(*dSpm*) cassettes are linked to the *iaaH* gene, the selection for a NAM^R phenotype would ensure the enrichment of unlinked jumps while simultaneously selecting against the *Ac* transposase, thereby generating stable insertions.

In the second step, the *trans*-activation of the *dSpm* borders was initiated from the primary transposition site to generate secondary transpositions by crossing the lines harboring the transposed *Ds* cassette with a *Spm* transposase containing line (Fig. 2). Resulting F2 or F3 seeds were selected on streptomycin containing medium to select for germinal *dSpm* excision events (seedlings with totally green phenotype on the selection medium). Re-integration of the excised *dSpm* element in such lines was analyzed by PCR.

Generation of the *Ds*(*dSpm*) starter lines

A total of 261 independent transgenic lines were generated in *A. thaliana* with the pPZP100: *Ds*(*dSpm*)

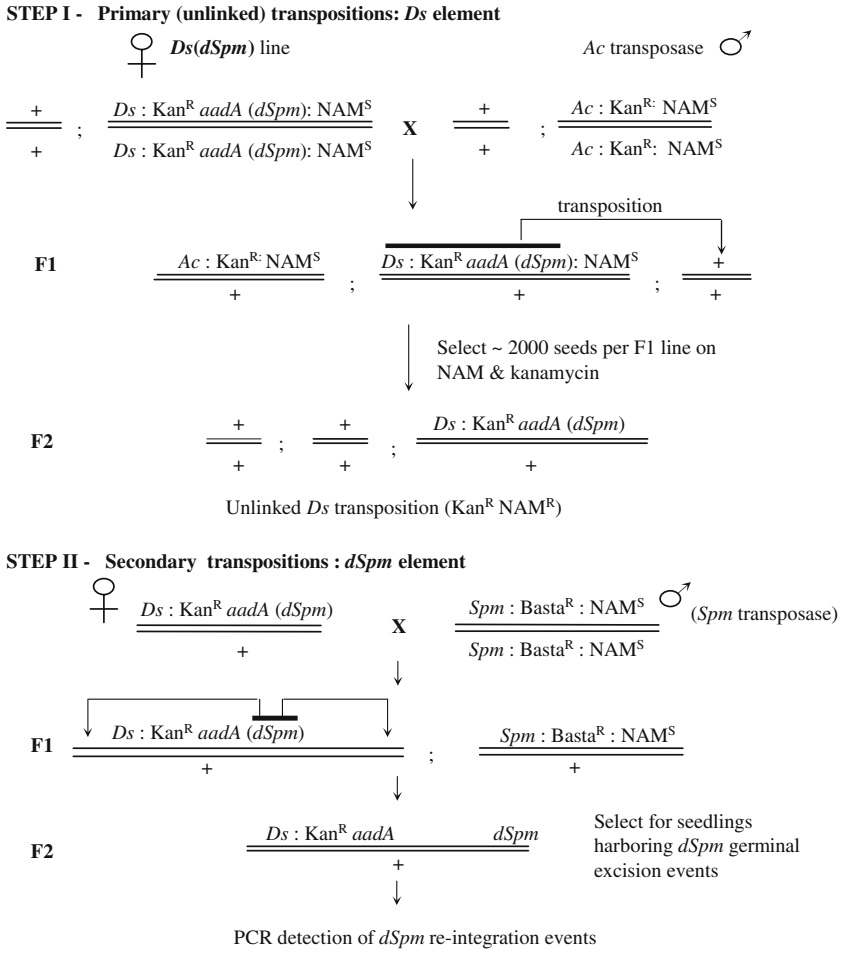

Fig. 2 Experimental strategy to achieve saturation mutagenesis or allele-specific tagging using the four-element system

construct (Fig. 1). On the basis of the segregation analysis of the F1 seedlings for kanamycin resistance, 65 lines were identified showing an inheritance pattern of T-DNA integrated at a single locus. These lines were further analyzed by Southern hybridization to determine the copy number of T-DNA using probes specific for the left and right border flanks. Five transgenic lines (~5% of the total lines analyzed) were found to harbor a single copy of the construct.

In order to map the location of the T-DNA insertions in the transgenic lines, the genomic region flanking the insertions was amplified using either TAIL-PCR or Genome Walking PCR. Primers specific to either the Ds3′ border (RB end) or LB region were used in different combinations with the arbitrary primers. The flanking sequences obtained were then subjected to BLASTN search (NCBI Genbank database) to map the insertions on the *Arabidopsis* genome. We could successfully map four of these single copy lines. The sites of integrations have been summarized in Table 1. Integrations were located on chromosome 2, 3 and 5. Homozygous lines of these events were developed and used in further crosses.

Selection for unlinked transpositions of the *Ds* element was based on the negative selection marker *iaaH*, which is also linked to the *Ac* transposase source (Fig. 2). Thus, it was preferable to use only those lines for crosses which had the *Ac* and the *Ds* construct on different chromosomes (i.e., they assorted independently of each other). As the physical location of the T-DNA in the sAc and the Ac2 lines was not known, this was assessed using a genetic approach. The F2 progeny of crosses between the four *Ds*(*dSpm*) lines and the two *Ac* lines were selected on kanamycin and NAM. Crosses showing $Kan^{S}NAM^{R}$ phenotype in 1/16th of their progeny indicated independently assorting loci. Thus the F1 progeny of crosses between 4E26 × sAc and 4E59 × Ac2 were taken up for further large-scale transposition experiments.

Functionality of *Ds* in the *Ds*(*dSpm*) cassette

Unlinked *Ds* transpositions were identified by screening ~2,000 seeds collected from individual F1 lines on germination medium supplemented with NAM (3.5 μM) and kanamycin (50 mg/l). The putative transposants (exhibiting the double resistant phenotype) were shifted to fresh selection plates to confirm the phenotype and subsequently transferred to soil for further characterization. We identified 19 F1 lines harboring unlinked *Ds* transpositions in their progeny from a total of 103 F1 lines generated from crosses between *Ds*(*dSpm*) line 4E59 and Ac2. The percentage of F2 seeds per F1 line exhibiting a $Kan^{R}NAM^{R}$ phenotype varied from 0.05 to 3.5%, of which 0.05% was the most frequent value exhibited by about three-fifths of the total number of F1 plants. F2 seeds from ~100 F1 plants from the cross between 4E26 and sAc were analyzed for transposition events. No *Ds* transpositions were observed in this case.

Table 1 Site of integration of the T-DNA in the single copy *Ds*(*dSpm*) lines and the transposed *Ds* element in $Kan^{R}NAM^{R}$ lines

Line name	Chromosome	Integration site in the chromosome (bp)	Locus tag	Gene description
Ds(*dSpm*) T-DNA harboring lines				
4E26	2	8376475		
4E59	3	18700469	At3G50380	Expressed protein
4E64	5	15119011		
4E90	3	15078		
Ds transposants ($K'an^{R}NAM^{R}$ lines)				
Ds59-1010	4	2330214		
Ds59-161	3	9547350		
Ds59-191	1	7966878	At1G22550	Proton-dependent oligopeptide transport (POT) family protein
Ds59-231	5	2127981	At5G06850	C2 domain containing protein
Ds59-251	3	9600476	At3G26230	Cytochrome P450 family protein
Ds59-453	4	1440164		
Ds59-581	2	7667721		
Ds59-711	5	22423409	At5G55240	Caleosin-related family protein/embryo-specific protein (putative)
Ds59-7656	5	302587	At5G01780	Oxidoreductase, 2OG-Fe(II) oxygenase family protein
Ds59-891	5	5744984	At5G17430	Ovule development protein (putative)
Ds59-981	5	17798126	At5G44170	Expressed protein

As linked *Ds* insertions leading to disruption of the *iaaH* gene can also result in a Kan^RNAM^R phenotype, the absence of *iaaH* gene in the 19 *Ds* transposants was tested by PCR. Two sets of *iaaH*-specific primers and primers from an internal *PROLIFERA* gene (as an internal control) were used for the PCR reaction (described by Martienssen and Springer, http://www.arabidopsis.org/home.html). Eighteen of the above Kan^RNAM^R lines (one each from the individual F1 lines) lacked the *iaaH* gene (data not shown) and were taken up for further characterization.

Localization of the *Ds* transpositions on *Arabidopsis* chromosomes

Genomic DNA from the double resistant (Kan^R-NAM^R) plants was subjected to Directional Genome Walking PCR. Sets of outwardly directed, nested primers specific to both ends of *Ds* were used in the PCR reactions. Flanking sequences could be amplified from 11 of the 18 Kan^RNAM^R plants using either the *Ds*3′ or the *Ds*5′ end-specific primers. For the other lines, either no amplification was observed or the overlapping border sequence was missing in the amplified fragments. Such inconclusive events were not analyzed further. Mapping the *Ds* insertions showed that these were distributed on all the five chromosomes (Fig. 3). The exact sites of integration in the genome are summarized in Table 1. Therefore, starting from a single *Ds* insertion (on chromosome 3) in the *Ds*(*dSpm*) line 4E59 insertions on all the five chromosomes of *A. thaliana* could be achieved. This demonstrated the functionality of the *Ds* cassette. These insertions would serve as new launch sites within the genome from where the *dSpm* can be *trans*-activated to generate further insertions.

Trans-activation of the *dSpm* element in the *Ds* transposants

dSpm transposition was initiated in the *Ds* transposed lines Ds59-1010, Ds59-191, Ds59-231 and Ds59-251 (Fig. 3) by crossing with the Spm404 line. As the allelic status of the *Ds* cassette in the transposed lines was not determined prior to crossing, the F1 seeds were selected for Kan^R to ensure the presence of the *Ds* construct. The Spm404 line is sensitive to kanamycin. F2 seeds from individual F1 lines were screened on streptomycin plates to analyze *dSpm* transpositions. On the basis of the presence of a completely green phenotype on selection medium, the germinal excision frequency was estimated to range between 0.005 and 0.03%. Majority of the F2 seedlings exhibited a variegated phenotype. Seedlings exhibiting green or large sectors of green (early excisions) phenotype were rescued on antibiotic-free medium and later transferred to soil for selfing. Approximately, 2,000 F3 seeds collected from individual plants were again subjected to streptomycin

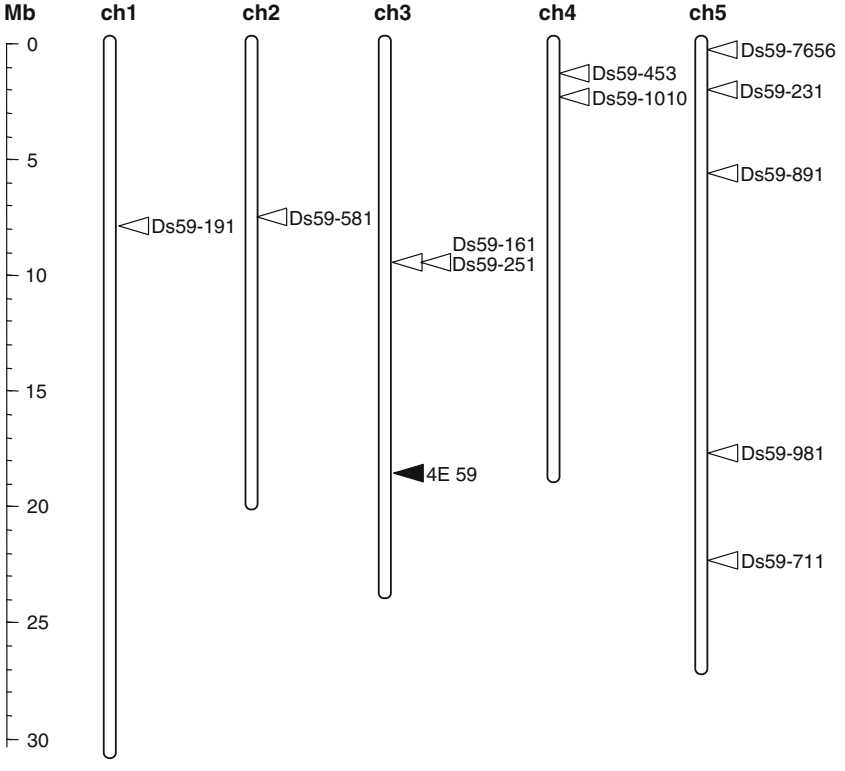

Fig. 3 Distribution of the *Ds*(*dSpm*) cassette in the *Arabidopsis* genome. The integrated map localizes the sites of integration of the *Ds* cassette in all the Kan^RNAM^R lines generated in the study. *Filled triangle* indicates the position of the *Ds* cassette in the parental line 4E59, while the *open triangles* represent the site of the transposed *Ds* element. The numbers represent the lines containing these insertions. The *scale bar* represents distance in megabase pairs (Mbp)

assay to select for germinal *dSpm* excisions. Seedlings exhibiting green phenotype were transferred to fresh selection plates for further characterization.

The re-insertion of the excised element was ascertained through PCR using *dSpm*-specific primers. To rule out the possibility of the *dSpm*-specific amplifications arising from the un-excised *dSpm* allele, the seedlings were first subjected to PCR with primers specific for the pre-excision allele (Fig. 1). From each F2 line, 20–30 green seedlings were analyzed and seedlings exhibiting no amplification products were analyzed further for the presence of the re-inserted *dSpm* element using primers designed from within the element (Fig. 1). Seedlings harboring the transposed *dSpm* element were transferred to soil. Genomic DNA isolated from these transposants was then subjected to TAIL-PCR to amplify the genomic region flanking the insertions using *dSpm*-specific primers designed from both the 5′ and 3′ border regions. We could localize *dSpm* insertions from three of the *Ds* transposant lines: Ds59–1010, Ds59-231 and Ds59-251. In case of the line Ds59-191, the rate of germinal *dSpm* excisions observed in the F3 generation was found to be very low for most of the F2 lines screened. Hence, this line was not used for further screenings. The distribution of the transposed *dSpm* elements on the *Arabidopsis* genome is shown in Fig. 4, while the exact sites of integrations are summarized in Table 2. From the Ds59-1010 line, 19 insertions could be selected of which three were present on the same chromosome. Seven *dSpm* insertions were characterized from the Ds59-251 line and three from the Ds59-231 line. Some of the lines had multiple insertions.

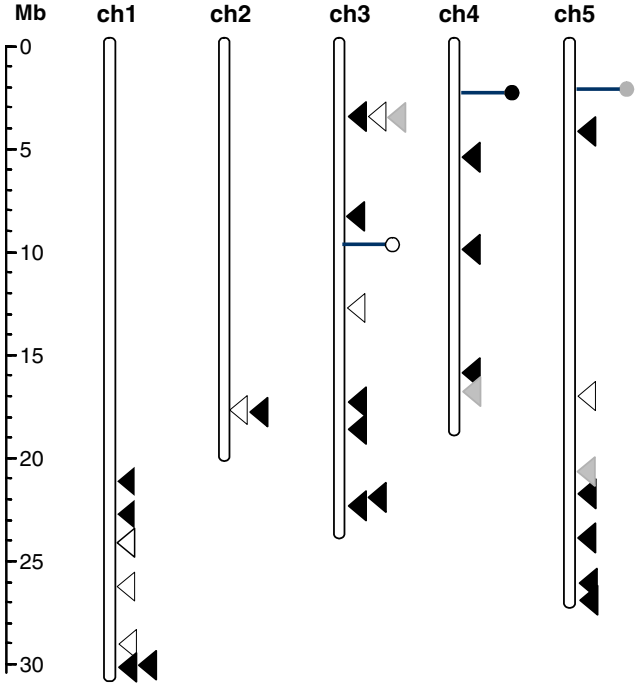

Fig. 4 Distribution of the *dSpm* element in the *Arabidopsis* genome. The *lollipop* structure indicates the site of integration of the *Ds*(*dSpm*) transposant and the *triangles* indicate transposed *dSpm* elements. The shade of the parental line and all the *dSpm* transpositions originating from it are identical. The *scale bar* represents distance in megabase pairs (Mbp)

Discussion

Our results show that the two transposon systems namely *Ac*/*Ds* and *Spm*/*dSpm* can work in conjunction and that the four-element system can functionally generate genome-wide independent insertions using a very limited number of starter lines. Starting from a single transgenic line containing the *Ds*(*dSpm*) construct, the transposon cassette could be distributed on all the different chromosomes of *Arabidopsis* thereby creating several new launch pads within the genome from where the internal *dSpm* could be activated *in-trans* to generate secondary transpositions. As the secondary transposition can be phenotypically selected, it saves the huge effort that is required in selecting new insertions.

From ~103 F1 lines generated by crossing a single *Ds*(*dSpm*) line 4E59 (insertion on chromosome 3) with *Ac* transposase, we could select for 18 unlinked *Ds* transpositions, of which 11 could be readily mapped on the genome. Two of these insertions were found on the same chromosome, loosely linked to the site of initial integration in the parental line. The other nine insertions were found distributed in the other four chromosomes. The frequency of unlinked *Ds* transpositions (Kan^RNAM^R progeny) in the progeny of F1 plants varied between 0.05 and 3.5%, of which 0.05% constituted the most frequent transposition frequency shown by ~60% of the F1 lines that were analyzed. In an earlier report (Sundaresan et al. 1995), a similar range (0.1–2.6%) of unlinked *Ds* transposition was reported wherein ~33% of the F1 lines exhibited 0.1% transposition frequency. Thus, in the present study the most common encountered transposition frequency of 0.05% was two-fold lower than the earlier report. One of the possible reasons for this drop could be the increased size (~8 kbp) of the *Ds* construct used in this study as compared to the construct used by Sundaresan et al. (1995). A *Ds* construct comparable to the size of the *Ds*(*dSpm*) construct has been used by Suzuki et al. (2001). However, the frequencies of unlinked transpositions have not been clearly described in this study. The size of the *Ds*(*dSpm*) cassette used in the present study can be reduced further by deploying smaller *Ds*

Table 2 Integration sites of the *dSpm* element in the *Arabidopsis* genome

Parental *Ds* transposant line (site of insertion)	Line	Chromosome	Integration site in the chromosome (bp)	Locus tag	Gene description
Ds59-1010 (chromosome 4)	A23.24.1	5	4333387	At5G13485	Non-LTR retrotransposon family (LINE)
	A23.24.2	5	26059708	–	–
	A23.24.3L	3	3355812	At3G10720	Pectinesterase, putative
	A23.24.3U	3	21983160	–	–
	A23.24.4	2	17535417	At2G41990	Expressed protein
	A23.24.9	3	17276165	–	
	A23.24.10	4	15934472	–	–
	A23.1.2	4	5416060	–	–
	A33.14.10	5	23883810	At5G59120	Subtilase family protein
	A33.14.20	5	21757786	At5G53510	Oligopeptide transporter OPT family protein
	A7.35.21	3	8288088	At3G23220	Member of the ERF (ethylene response factor) subfamily B-3 of ERF/AP2 transcription factor family
	A7.25.10	1	30063018	AT1G79915	Expressed protein
	A7.24.1	1	30177790	–	–
	A7.24.4	3	22386381	AT3G60540	Sec61beta family protein
	A7.24.7L	5	26758132	AT5G66970	Signal recognition particle-related/SRP-related
	A7.24.7U	3	18696673	At3G50376	Hypothetical protein
	A33.16.1	4	9851084	–	–
	A33.16.11	1	22829169	At1G61800	Glucose-6-phosphate/phosphate translocator, putative
	A5.1.28	1	21174015	AT1G56510	Disease resistance protein (TIR-NBS-LRR class), putative
Ds59-251 (chromosome 3)	E13.7.4	5	17052551	–	–
	E13.3.22L	1	26302365	AT1G69850	Nitrate transporter (NTL1)
	E13.3.22U	3	3492979	AT3G11150	Expressed protein
	E13.7.1	2	17551736	AT2G42040	Expressed protein
	E13.7.3	3	12666947	AT3G31023	Copia-like retrotransposon family
	E13.7.6a	1	24013110	–	–
	E13.7.6b	1	28975697	–	–
Ds59-231 (chromosome 5)	D12.4.7	4	16816586	At4G35370	Transducin family protein/WD-40 repeat family
	D12.4.3	5	20747677	At5G50950	Fumarate hydratase, putative
	D12.4.5	3	3492960	At3G11150	Expressed protein

and *dSpm* borders. Significant variation in the frequency of *Ds* transpositions from different single copy T-DNA donor sites has been observed (Smith et al. 1996), suggesting that chromatin structure influences transposition frequency. This can also be a possible reason for the lower frequency of transpositions observed in this study since the results presented are from a single T-DNA insertion.

After the initial unlinked spread of the *Ds*(*dSpm*) construct, the internally placed *dSpm* element was activated *in-trans* in four of the *Ds* transposant lines. These four lines gave a germinal frequency ranging between 0.005 and 0.03%. The frequency, although low, is comparable to the frequency range of 0.0009–4.7% reported by Tissier et al. (1999) wherein most of the plants exhibited germinal excision frequency of less than 1%. The frequencies recorded in this study are in the lower range of the earlier report which could be due to the fewer number of plants that were analyzed.

Of the 29 *dSpm* insertions studied, six were linked and the rest were unlinked transposition events. The *Spm* element has been reported to transpose to linked sites at a frequency of 60% in maize and around 45% in tobacco (Cardon et al. 1993). An earlier study in *Arabidopsis* (Tissier et al. 1999) has recorded more than 80% of the transpositions to be linked. However,

our study shows the opposite; around 80% of the *dSpm* transpositions belong to unlinked sites. These insertions were spread on all the five chromosomes of *Arabidopsis*. Hence, even with a minimum of three *Ds*(*dSpm*) launch pads, *dSpm* can be suitably *trans*-activated to generate a large number of insertions distributed throughout the genome. *dSpm* elements from the rest of the *Ds* launch sites can be similarly *trans*-activated to generate more insertions for saturating the genome.

Currently, saturation mutagenesis using transposons is achieved by either unlinked dispersal of the transposon construct from a few starter lines or localized mutagenesis using several starter lines. The first strategy of genome-wide coverage using unlinked transpositions is an arduous task since the frequency of such transpositions is low (Sundaresan et al. 1995; Tissier et al. 1999). The second strategy is applicable to model systems like *Arabidopsis* but may not be feasible in plants less amenable to genetic transformation. Many crop plants belong to the latter category. The difficulty in using the second strategy is further compounded by the need to identify single-copy starter lines. In such situations a four-element system can be advantageous since a minimum of two transgenic lines containing the *Ds*(*dSpm*) construct, apart from the lines expressing the specific transposases, will be sufficient to initiate a saturation mutagenesis program in a species. Although, in a two-element system, the *Ds* or the *dSpm* insertions resulting from the initial transpositions can be re-activated to generate further insertions, the absence of a visual excision marker makes the task of identifying new transposition events tedious. Moreover, many of the launch sites are rendered transposition-incompetent due to the perturbations in the borders during the initial transposition (Ito et al. 2002). In the present strategy this difficulty is overcome by activating a second transposon carrying its own borders and independent transposition markers.

Although the present study has conclusively shown the functionality of the *Ds*(*dSpm*)-based four-element system for gene tagging, the system can be further modified to optimize its functionality in *Arabidopsis* as well as in other plant species. The kanamycin resistance conferring *nptII* gene, currently placed outside the *dSpm* cassette (Fig. 1), can be cloned within the *dSpm* borders, thereby allowing the use of *nptII* both as a transformation as well as an integration marker for *Ds* and *dSpm* transpositions. With this type of construct the $Kan^{R}NAM^{R}$ phenotype would help in selecting unlinked *Ds* transpositions, while the $Strep^{R}Kan^{R}$ plants would be indicative of *dSpm* transpositions. The use of a herbicide resistance conferring marker (viz., *bar* gene) will allow field selection of transpositions. Incorporation of a negative selection marker more effective than *iaaH* would further facilitate the screening process. We found *iaaH* to be a very cumbersome selection marker since it does not kill the plants carrying it. The selection is based on morphological changes. Using proficient pro-herbicide to herbicide conversion based negative selection marker like SUI (Tissier et al 1999; Marsch-Martinez et al. 2002; Greco et al. 2004) would be a better choice. The use of markers like SUI would also allow field level selections and thus extend the use of the four-element system to crop species. Fluorescence-based non-destructive selection markers like GFP (green fluorescent protein) and RFP (red fluorescent protein) have been shown to work efficiently in *Arabidopsis* (Kumar et al. 2005). These can also be used to enhance the screening efficiency. Additionally, *in-cis* expression of the *Ac* Tpase (Tissier et al. 1999), especially under an inducible promoter (Nishal et al. 2005), can reduce the number of generations involved and also increase the frequency of selecting unlinked jumps. An inducible promoter would limit the extent of the transposase activity, thus reducing the background mutations which result from the continued presence of the transposase.

A further improvement in the *Ds*(*dSpm*) construct could be the use of a *dSpm*(*Ds*) construct in which *dSpm* is used for unlinked transpositions and *Ds* is used subsequently for linked transpositions for saturation mutagenesis. The present study has conclusively shown that *dSpm* transposes to unlinked sites at a much higher frequency than what has been reported for the *Ds* element (Ito et al. 2002). The *Ac*/*Ds* transposable elements have also been shown to be more versatile in terms of frequency of transposition in evolutionary very diverse species (Ramachandran and Sundaresan 2001). Novel transposon elements discovered in specific species (Ramachandran and Sundaresan 2001) could be used in conjunction with *Ds* for developing four-element systems tailored for specific plant species.

Although a number of strategies for mutagenesis are available, systems based on transposable elements like the four-element system described in this study will remain very cost-effective and versatile methods for saturation mutagenesis and site-specific mutagenesis.

Acknowledgments We thank Dr. V. Sundaresan for the gift of constructs pWS32 and pAJ6, the low copy vector pMuCBs and the transposase line Ac2; Dr. R. Schmidt for providing the transposase line sAc; Dr. Monika Frey for the transposase line Spm404; Dr. J.D.G. Jones for the construct pSLJ532 harboring the 35S-*spec*-pA cassette and Dr. R. Raina for the construct pEMBL118-*dSpm* used for generating the *dSpm* borders used in

the present study. We are grateful to Dr. Sanjay Phogat for useful discussions and suggestions. We thank Shankar for technical assistance in screening for transpositions and maintenance of the growth rooms and Prasad Rao and UmaShanker for assistance in infiltration experiments. This work was supported by grant-in-aid from CSIR (Council of Scientific and Industrial Research) under the NMITLI (New Millennium Indian Technology Leadership Initiative) programme. Priya Panjabi was supported by a CSIR research fellowship and a post-doctoral research fellowship by MAHYCO (Maharashtra Hybrid Seeds Company Ltd.).

References

Altmann I, Schmidt R, Willmitzer L (1992) Establishment of a gene tagging system in *Arabidopsis thaliana* based on the maize transposable element *Ac*. Theor Appl Genet 84:371–383

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

Bancroft I, Bhatt AM, Sjodin C, Scofield S, Jones JDG, Dean C (1992) Development of an efficient two-element transposon tagging system in *Arabidopsis thaliana.* Mol Gen Genomics 233:449–461

Bechtold N, Ellis J, Pelletier G (1993) *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C R Acad Sci Paris 316:1194–1199

Cardon GH, Frey M, Saedler H, Gierl A (1993) Definition and characterization of an artificial *En*/*Spm*-based transposon tagging system in transgenic tobacco. Plant Mol Biol 23:157–178

Chen S, Hofius D, Sonnewald U, Bornke F (2003) Temporal and spatial control of gene silencing in transgenic plants by inducible expression of double-stranded RNA. Plant J 36:731–740

Greco R, Ouwerkerk PBF, Taal AJC, Sallaud C, Guiderdoni E, Meijer AH, Hoge JHC, Pereira A (2004) Transcription and somatic transposition of the maize *En*/*Spm* transposon system in rice. Mol Gen Genomics 270:514–523

Guo H-S, Fei J-F, Xie Q, Chua N-H (2003) A chemical-regulated inducible RNAi system in plants. Plant J 34:383–392

Hajdukiewicz P, Svab Z, Maliga P (1994) The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. Plant Mol Biol 25:989–994

Hehl R, Baker B (1989) Induced transposition of *Ds* by a stable *Ac* in crosses of transgenic tobacco plants. Mol Gen Genomics 217:53–59

Ito T, Motohashi R, Kuromori T, Mizukado S, Sakurai T, Kanahara H, Seki M, Shinozaki K (2002) A new resource of locally transposed *Dissociation* elements for screening gene-knockout lines *in silico* on the *Arabidopsis* genome. Plant Physiol 129:1695–1699

Kiegle E, Moore CA, Haseloff J, Tester MA, Knight MR (2000) Cell-type-specific calcium responses to drought, salt and cold in the *Arabidopsis* root. Plant J 23:267–278

Klimyuk VI, Carroll BJ, Thomas CM, Jones JDG (1993) Alkali treatment for rapid preparation of plant material for reliable PCR analysis. Plant J 3:493–494

Kumar CS, Wing RA, Sundaresan V (2005) Efficient insertional mutagenesis in rice using the maize *En*/*Spm* elements. Plant J 44:879–892

Li X, Song Y, Century K, Straight S, Ronald P, Dong X, Lassner M, Zhang Y (2001) A fast neutron deletion mutagenesis-based reverse genetics system for plants. Plant J 27:235–242

Liu YG, Whittier RF (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

Liu YG, Mitsukawa N, Oosumi T, Whittier RF (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J 8:457–463

Long D, Swinburne J, Martin M, Wilson K, Sundberg E, Lee K, Coupland G (1993) Analysis of the frequency of inheritance of transposed *Ds* elements in *Arabidopsis* after activation by a CaMV 35 S promoter fusion to the *Ac* transposase gene. Mol Gen Genomics 241:627–636

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

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

Masterson RV, Furtek DB, Grevelding C, Schell J (1989) A maize *Ds* transposable element containing a dihydrofolate reductase gene transposes in *Nicotiana tabacum* and *Arabidopsis thaliana*. Mol Gen Genomics 219:461–466

McCallum CM, Comai L, Greene EA, Henikoff S (2000) Targeting induced local lesions in genomes (TILLING) for plant functional genomics. Plant Physiol 123:439–442

Mishra RN, Singla-Pareek SL, Nair S, Sopory SK, Reddy MK (2002) Directional genome walking using PCR. Biotechniques 33:830–834

Nishal B, Tantikanjana T, 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 WC, Kumaran M, Sundaresan V (1999) Analysis of flanking sequences from *Dissociation* insertion lines: a database for reverse genetics in *Arabidopsis*. Plant Cell 11:2263–2270

Phogat S, Burma PK, Pental D (2000) A four-element based transposon system for allele specific tagging in plants—theoretical considerations. J. Biosci 25(1):57–63

Ramachandran S, Sundaresan V (2001) Transposons as tools for functional genomics. Plant Physiol Biochem 39:243–252

Rogers SO, Bendich AJ (1994) Extraction of total cellular DNA from plants, algae and fungi. In: Gelvin SB, Schilperoot RA (eds) Plant Mol Biol Manual. Kluwer Academic Publishers, Dordrecht, pp D:1–D:8

Slade AJ, Fuerstenberg SI, Loeffler D, Steine MN, Facciotti D (2005) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. Nat Biotechnol 23:75–81

Smith DL, Fedoroff NV (1995) *LRP1*, a gene expressed in lateral and adventitious root primordia of *Arabidopsis*. Plant Cell 7:735–745

Smith D, Yanai Y, Liu Y-G, Ishiguro S, Okada K, Shibata D, Whittier RF, Fedoroff NV (1996) Characterization and mapping of Ds-GUS T-DNA lines for targeted insertional mutagenesis. Plant J 10:721–732

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

Suzuki Y, Uemura S, Saito Y, Murofushi N, Schmitz G, Theres K, Yamaguchi I (2001) A novel transposon tagging element for obtaining gain-of-function mutants based on a self-stabilizing *Ac* derivative. Plant Mol Biol 45:123–131

Swinburne J, Balcells L, Scofield SR, Jonathan DGJ, Coupland G (1992) Elevated levels of *Activator* transposase mRNA are associated with high frequencies of *Dissociation* excision in *Arabidopsis*. Plant Cell 4:583–595

Till BJ, Reynolds SH, Greene EA, Codomo CA, Enns LC, Johnson JE, Burtner C, Odden AR, Young K, Taylor NE, Henikoff JG, Comai L, Henikoff S (2003) Large-scale discovery of induced point mutations with high throughput TILLING. Genome Res 13:524–530

Tissier AF, Marillonnet S, Klimyuk V, Patel K, Torres MA, Murphy G, Jones JDG (1999) Multiple independent defective *suppressor-mutator* transposon insertions in *Arabidopsis*: a tool for functional genomics. Plant Cell 11:1841–1852

Tsugeki R, Kochieva EZ, Fedoroff NV (1996) A transposon insertion in the *Arabidopsis SSR16* gene causes an embryo-defective lethal mutation. Plant J 10:479–489

Wilson K, Long D, Swinburne J, Coupland G (1996) A *Dissociation* insertion causes a semidominant mutation that increases expression of *TINY*, an *Arabidopsis* gene related to *APETALA2*. Plant Cell 8(4):659–671