

Isolation of 151 Mutants that Have Developmental Defects from T-DNA Tagging

Ji Hoon Ahn^{1,3,*}, Joonki Kim^{1,3}, Seong Jeon Yoo¹, So Yeon Yoo¹, Hyungmin Roh², Jun-Hyuk Choi^{2,4}, Mi Suk Choi^{1,4}, Kyung Sook Chung^{1,4}, Eun Ju Han^{2,4}, Sung Myun Hong^{1,4}, Sung Hye Jung^{1,4}, Hyo Jin Kang^{2,4}, Bo Kyung Kim^{2,4}, Mi Duk Kim^{2,4}, Youn Kyung Kim^{2,4}, Yun Hee Kim^{2,4}, Hanna Lee^{1,4}, Soo Hyun Park^{1,4}, Jee Hoon Yang^{1,4}, Jung Won Yang^{2,4}, Dong-Hun Yoo^{2,4}, Seung Kwan Yoo^{1,4} and Jong Seob Lee^{2,*}

¹ Plant Signaling Network Research Center, School of Life Sciences and Biotechnology, Korea University, Seoul 136-701, Korea

² School of Biological Sciences, Seoul National University, Seoul 152-742, Korea

In order to understand the mechanisms underlying plant development, a necessary first step involves the elucidation of the functions of the genes, via the analysis of mutants that exhibit developmental defects. In this study, an activation tagging mutant library harboring 80,650 independent *Arabidopsis* transformants was generated in order to screen for developmental mutants. A total of 129 mutants manifesting dominant developmental abnormalities were isolated, and their T-DNA insertion loci were mapped. The activation of one or more genes adjacent to a T-DNA insertion locus was confirmed in eight dominant mutants. A gene adjacent to the right border was usually activated by the 35S enhancers. Interestingly, the transcriptional activation of multiple genes within a broad range was observed in one of the mutants, which raises the possibility that activation by the 35S enhancers was not limited strictly to a single gene. In order to gain a better understanding of sexual reproduction in higher plants, we isolated 22 mutants exhibiting defects in female gametophyte development, and determined their T-DNA insertion loci. We propose that this mutant population may prove useful in the further determination of the functions of genes that play important roles in plant development.

Keywords: Activation tagging — *Arabidopsis* — Developmental mutants — Female gametophytic mutant — Plant development.

Abbreviations: BAC, bacterial artificial chromosome; LB, left border; RB, right border; RT-PCR, reverse transcription-PCR; TAIL-PCR, thermal asymmetric interlaced-PCR.

Introduction

After the completion of the *Arabidopsis* genome project (Arabidopsis Genome Initiative 2000), one of the goals has been the elucidation of the functions of all the genes in the *Arabidopsis* genome, and the acquisition of a greater understanding of the mechanisms underlying the development of the plant (Chory et al. 2000). Among the variety of approaches to the achievement of this goal, the most straightforward way to identify gene function is to isolate a mutant with a specific, inactivated gene, to analyze its phenotype and then to infer the role of the gene under normal conditions. To date, many groups have generated *Arabidopsis* loss-of-function mutant populations on a large scale, and have identified genes that perform important functions in plant development (Ostergaard and Yanofsky 2004).

The classical loss-of-function approach tends to be limiting. This is because the absence of a visible phenotype can restrict the identification of such mutants. This problem becomes more salient when two or more redundant genes perform certain functions (Liljegren et al. 2000) or when a mutation induces early embryonic lethality (Naoui and Hashimoto 2004). An additional hindrance is that not all genes within the *Arabidopsis* genome have thus far been tagged, despite the tremendous efforts made to generate large-scale mutant populations. These unsaturated genetic resources prove problematic when a class of genes, which are not predicted by conventional methods but play important roles in plant development, are identified at

³These authors contributed equally to this work.

⁴These authors contributed equally to this work and are listed in alphabetical order.

*Corresponding authors: Jong Seob Lee, E-mail, jongslee@plaza.snu.ac.kr; Fax, +82-2-872-1993; Ji Hoon Ahn, E-mail, jahn@korea.ac.kr; Fax, +82-2-927-9028.

a later date, as occurred in the discovery of microRNAs (Moss 2002).

One of the most frequently applied strategies for the circumvention of the problems inherent in the classical loss-of-function approach is activation tagging (Weigel et al. 2000), which screens for dominant gain-of-function phenotypes. In addition to its conventional advantages, activation tagging has proven especially useful in the isolation of genes that play roles in plant development. For example, an unforeseen class of genes, i.e. microRNAs, which play fundamental roles in plant development, was identified via activation tagging (Palatnik et al. 2003). Activation tagging is also useful for the discovery of the functions of genes that have dual functions in both early embryonic and adult stages (Marsch-Martinez et al. 2002). In the *Arabidopsis* genome, approximately 1,200 genes have been theorized to be essential in early development (Vizir et al. 1994). As these essential genes often function in adult stages, viable mutants of such essential genes can be isolated via activation tagging. Taken together, although the generation of an activation tagging population requires a great deal of time and effort, this population may constitute a comprehensive genetic resource for the study of gene functions in plant development (Ichikawa et al. 2003, Nakazawa et al. 2003, Tani et al. 2004, Schneider et al. 2005, Seki et al. 2005, Perrella et al. 2006).

The female gametophyte, the structure which generates the egg cell and the central cell (Drews and Yadegari 2002), performs critical functions in sexual reproduction, as sexual reproduction provides a highly selective evolutionary advantage to plants (Charlesworth and Charlesworth 1979). The female gametophyte is also the structure that originates the seed's embryo and endosperm. The female gametophyte mediates a number of reproductive processes, including pollen tube guidance, fertilization and the induction of seed development (Drews and Yadegari 2002). Although several important genes, including *DEMETER* (*DME*) and *FERTILIZATION INDEPENDENT ENDOSPERM* (*FIE*), have been identified as important regulators of female gametophyte development (Yadegari and Drews 2004), more genes must be isolated in order to determine the precise mechanism underlying the development of the female gametophyte.

In this study, 80,650 activation-tagged mutants were generated in order to screen for developmental mutants. We isolated 151 mutants that showed developmental abnormalities, including 22 mutants that were defective in female gametophyte development, and determined the T-DNA insertion sites of 5,361 of the mutants. We confirmed the activation of a gene(s) adjacent to the 35S enhancers in eight dominant mutants. The information

was integrated in a database in order to study plant development systematically. We propose that this mutant population might prove useful in the isolation of genes relevant to plant development.

Results

Generation of an activation tagging mutant library

A total of 80,650 activation-tagged plants were generated in this study. The estimated number of T-DNAs in our mutant library was 1.2 copies per line (Supplementary Fig. S1). These plants were screened to determine whether or not they manifested developmental defects in the T₁ generation. During phenotypical screening, the co-segregation of dominant phenotypes with BASTA resistance was monitored. Among 80,650 mutants, seeds from 45,116 individual lines were ultimately deposited in the stock center of the Crop Functional Genomics Center (<http://cfgc.snu.ac.kr/english/index.html>) in Korea.

Amplification of sequences flanking T-DNA insertion loci and sequence determination

Approximately 62% of the lines processed in this study generated visible PCR products from thermal asymmetric interlaced PCR (TAIL-PCR), using the DEG1 primer (Fig. 1A). Multiple PCR products were observed in 15% of mutants in which TAIL-PCR generated PCR products, thereby suggesting that these mutants may harbor multiple T-DNA insertions. For those that did not yield amplified products, TAIL-PCR using the DEG2 primer was conducted; however, the majority of these again failed to generate PCR products. Only 44% of the starting mutants generated readable sequencing results. However, the pSKI015 vector sequences alone were retrieved from a small number of mutants, which may be the result of an incomplete removal of the oligonucleotides used in TAIL-PCR. After discarding these false positives, we determined that the ratio of identifying T-DNA insertion loci was approximately 41%.

We obtained information regarding multiple insertion loci from a single sequencing run which produced overlapping sequencing peaks (Fig. 1B), as was previously shown (Sessions et al. 2002). These overlapping peaks probably appeared because we conducted TAIL-PCR without purifying individual amplified bands. These mutants that generated overlapping sequencing peaks were considered to harbor more than one insertion, which is consistent with the observation that 14% of mutants in our library had been predicted to harbor multiple insertions (Supplementary Table S1). For example, in the RN10846 mutant, two products, with sizes of 160 and 450 bp (designated fragments A and B, respectively), were amplified. The amplitudes of the peaks from fragment A were higher, possibly due to the fact that small bands were more

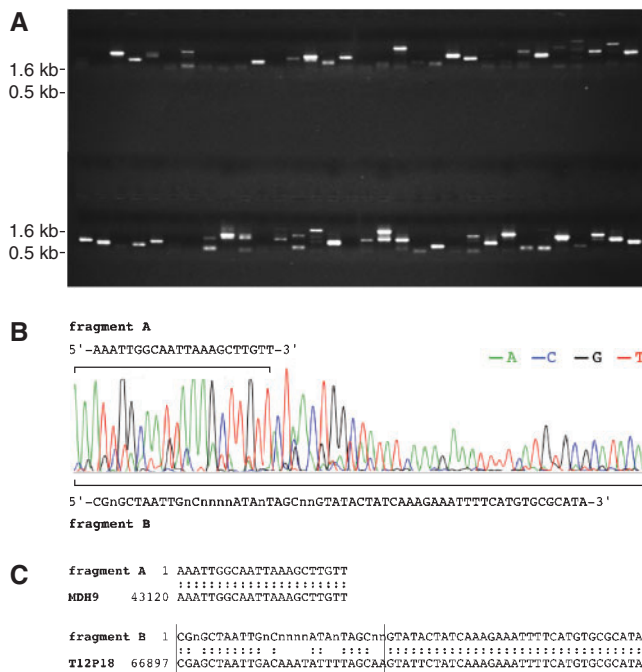


Fig. 1 Sequence analysis of the TAIL-PCR products. (A) Separation of products of TAIL-PCR using the DEG1 primer on a 1% agarose gel. Size markers are shown on the left, with band sizes. (B) Part of sequencing results in the RN10846 mutant. Two different products were amplified from the RN10846 mutant, fragment A (160 bp) and fragment B (450 bp). Peaks from fragment A masked the lower peaks from fragment B until the sequencing reaction from fragment A was terminated. Individual color-coded peaks were converted to the corresponding sequences. An ambiguous sequence from fragment B is shown by n. (C) Comparison of sequences of fragments A and B with the genomic sequences of *Arabidopsis*. Sequences of fragment B in an open box indicate the sequences that were manually read from lower amplitude peaks masked by strong peaks from fragment A. Note that these manually read sequences also hit the BAC clone T12P18 sequences. Identical sequences are indicated by colons.

effectively amplified, and showed predominating signals at the beginning of a sequencing run (Fig. 1B). These dominant peaks yielded information regarding an insertion locus in the MDH9 bacterial artificial chromosome (BAC) clone. After the termination of a sequencing reaction with fragment A, peaks from fragment B were apparent, albeit weak, such that another insertion locus in the T12P18 BAC clone was identified. Multiple insertions from these chimeric sequences were further confirmed via the manual reading of minor peaks (5'-CGnGCTAATTGnCnnnnATAnTAGCnnGTATAC-3') which had been masked by the prominent signals from fragment A (Fig. 1C).

Mapping T-DNA insertion loci

The T-DNA insertion loci of 5,361 mutants were determined. Among 5,361 lines, the insertion sites of 4,028

randomly selected insertions were mapped on *Arabidopsis* chromosomes (Fig. 2). Each of the chromosomes harbored 1,073, 651, 793, 570 and 941 insertions, respectively, thereby indicating a similar frequency of insertion into each of the chromosomes, with the average insertion frequency being one per 113 kb. The genes adjacent to the T-DNA insertion sites were classified in accordance with the guidelines established in the Functional Catalogue of the Munich Information Center for Protein Sequences (MIPS) (Supplementary Table S2) (Ruepp et al. 2004).

Identification of mutants that exhibited developmental defects

Of 80,650 mutants, 431 mutants that showed dominant developmental abnormalities were isolated in the T₁ generation (Table 1). This indicates that approximately 0.5% of our activation-tagged mutants exhibited dominant phenotypes, slightly lower than the average frequency at which gain-of-function mutants are detected (Weigel et al. 2000). This low frequency may be attributable to the silencing of the 35S enhancers (Chalfun-Junior et al. 2003). Among 431 mutants, predicted T-DNA insertion loci were determined in 129 of the mutants (Supplementary Table S3).

Confirmation of activation of a gene adjacent to a T-DNA insertion

In order to identify a gene responsible for the dominant phenotype, the expression levels of genes adjacent to a T-DNA were determined. Eight mutants in which an up-regulated gene(s) was identified are shown in Fig. 3. Phenotypes observed in these mutants were segregated as a dominant trait into the T₂ generation (Supplementary Table S4), which suggests that the phenotypes were probably the result of an up-regulation of a gene in each line. To facilitate discussion, genes closest to the left border (LB) and the right border (RB) are indicated by *Gene L1* and *Gene R1*, respectively. The second closest genes to the RB are similarly designated as *Gene R2*.

The S17-026 mutant showed increased apical dominance and epinastic leaves (Fig. 3A), characteristics which are similar to those of mutants defective in auxin biosynthesis (Zhao et al. 2001). The expression level of *Gene R1* was enhanced in this mutant, whereas *Gene L1* expression was not affected, suggesting that *Gene R1* is responsible for its phenotype. Accelerated flowering was observed in the K35731 mutant (Fig. 3B). The K35731 mutant flowered with 6.7 and 6.8 leaves under long- and short-day conditions, respectively, which suggests that this early flowering phenotype occurred independently of the photoperiod. This mutant also showed up-regulation of *Gene R1*. Early flowering phenotype and terminal flower formation were observed in the KSNU3 mutant (Fig. 3C). This mutant also manifested the same flowering time under

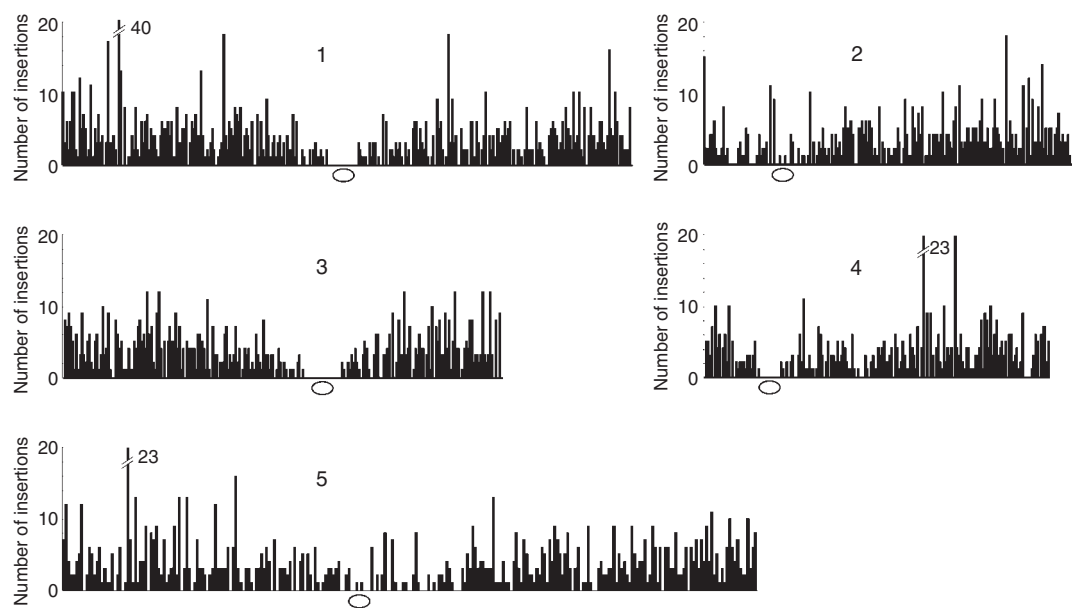


Fig. 2 Distribution of predicted T-DNA insertions in the *Arabidopsis* genome. The number of insertions per 100 kb interval was plotted as peaks on each chromosome. The locations harboring >20 insertions are indicated by numbers. Open ovals indicate centromere positions.

Table 1 Summary of mutants that showed dominant developmental abnormalities

Leaf		Flower		Flowering time		Others	
Leaf shape	82	Flower size	8	No flowering	5	Dwarfism	29
Leaf color	22	Determinacy	14	Early flowering	48	Bushy	65
Petiole length	23	Floral organ	8	Late flowering	35	Miscellaneous	73
Trichome formation	4	Silique	15				

both long- and short-day conditions (4.5 vs. 4.7 leaves), thereby suggesting that the activation of *Gene R1* in the KSNU3 mutant resulted in photoperiod insensitivity. Serrated leaves and retarded growth were observed in the S6-003 line, in which *Gene R1* up-regulation was detected (Fig. 3D). In addition to the leaf shape, spongy parenchymal cell layers were absent in the S6-003 mutant (data not shown), which indicates that the up-regulation of *Gene R1* primarily affects leaf development. A long petiole and hyponasty were observed in the RN13443 mutant (Fig. 3E), which represents a phenotype typically observed during shade avoidance (Pierik et al. 2004), with the exception of stem elongation. *Gene R1* expression was up-regulated, whereas the expression of *Gene L1* was not affected by the 35S enhancers. Short stature and clustered leaves were the main phenotypes observed in the S10-004 mutant (Fig. 3F), which phenocopied the brassinosteroid mutants (Friedrichsen et al. 2000). This distinct phenotype was rescued via brassinolide treatment (Supplementary Fig. S1). Furthermore, this mutant proved to be sensitive

to exogenous brassinazole, thereby suggesting that an unidentified gene in the S10-004 mutant dominantly inhibits brassinosteroid signaling. Interestingly, expression of *Genes L1* and *R1* was up-regulated in this mutant. Severe fasciation in the stem and clustered flowers was observed in the S27-001 mutant (Fig. 3G), which was a phenocopy of *Meristem enlargement1 (MEN1)*, a dominant mutant of *miR166a* (Kim et al. 2005). Indeed, we determined that the expression of *miR166a* was up-regulated in this mutant (data not shown). Interestingly, *Gene L1* expression was also enhanced in the S27-001 mutant. This result, along with those obtained with the S10-004 mutant (Fig. 3F), raised the possibility that transcriptional enhancement by the 35S enhancers may not be strictly limited to *Gene R1*. Shortened petioles were observed in the RN23916 mutant (Fig. 3H). This phenotype became more apparent in older leaves, suggesting that petiole elongation is inhibited or that leaf blade formation occurs at the petiole region late in leaf development. Interestingly, *Gene L1* and *R1* expression appeared to be unaffected; however, the expression of *Gene*

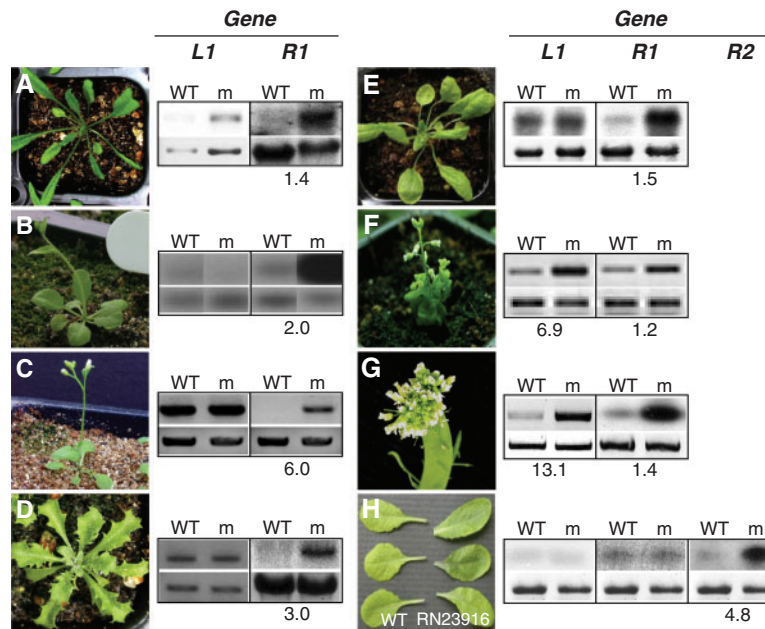


Fig. 3 Expression analysis of genes adjacent to a T-DNA insertion. RT-PCR or Northern hybridization were conducted in order to identify an up-regulated gene in each of the mutants. The upper panels indicate the expression levels of genes adjacent to a T-DNA. Either *UBQ10* or an actin gene was used as an internal control, and are shown in the bottom panels. Numbers denote the distance (kb) from the 35S enhancers to a gene whose expression is up-regulated. All of these mutants were hemizygous when pictures were obtained. m, mutant; WT, wild-type plants. (A–E) Expression analysis revealed that the transcription level of *Gene R1* was increased in these mutants; S17-026 (A), K35731 (B), KSNU3 (C), S6-003 (D) and RN13443 (E). (F–G) Transcription levels of *Gene L1* and *Gene R1* were enhanced in S10-004 (F) and S27-001 (G) mutants. (H) Up-regulation of *Gene R2*, but not *Gene R1*, was detected in the RN23916 mutant.

R2 was enhanced by the 35S enhancers, thereby suggesting that the 35S enhancers may remotely activate the expression of a gene without affecting *Gene R1* expression.

Collectively, of eight analyzed mutants, seven mutants showed *Gene R1* up-regulation. This suggests that *Gene R1* is generally responsible for the dominant phenotype observed in an activation tagging mutant. To confirm this notion, we expressed *Gene R1* of S17-026 and K35731 mutants under the control of the 35S promoter. Indeed, the overexpression of *Gene R1* of S17-026 and K35731 mutants was sufficient to recapitulate the original phenotypes (Supplementary Fig. S2), suggesting that *Gene R1* is responsible for the dominant phenotype.

Transcriptional activation of multiple genes by the 35S enhancers

In order to test the range of transcriptional activation by the 35S enhancers, the genomic responses of activation were monitored via microarray analysis. A mutant that showed dominant early flowering, K35731 (Fig. 3C), was selected for this experiment. The results showed that the RNA abundance of 1,273 genes was elevated by >2-fold in the K35731 mutant. Interestingly, of 1,273 genes, 32 genes were localized in a 984 kb region which contained 244 genes adjacent to the RB of a T-DNA (Fig. 4A). It appeared that

up-regulated and down-regulated genes appeared to distribute randomly outside of this 984 kb region. A close examination indicated that among 244 genes within the 984 kb region, 201 genes were up-regulated to variable extents, whereas only six genes were down-regulated by >2-fold.

The up-regulation of genes that located remotely from a T-DNA insertion site within the 984 kb region was verified via reverse transcription-PCR (RT-PCR) (Fig. 4B). We monitored the expression levels of seven randomly selected genes within the 984 kb region. Among these seven genes, six displayed increased mRNA abundance, although the extent of up-regulation varied substantially. In particular, increased expression levels of At2g43610 and At2g43100, which are located 720 and 890 kb away from a T-DNA insertion, respectively, were observed. This indicates that the 35S enhancers might affect gene expression within a broad region. This further suggests that any gene within the affected region, not simply a gene adjacent to the RB, may be a candidate for the dominant phenotype observed in an activation tagging line.

Identification of female gametophytic mutants

A total of 111 mutants that manifested aberrant embryo sac formation during gametophytic development

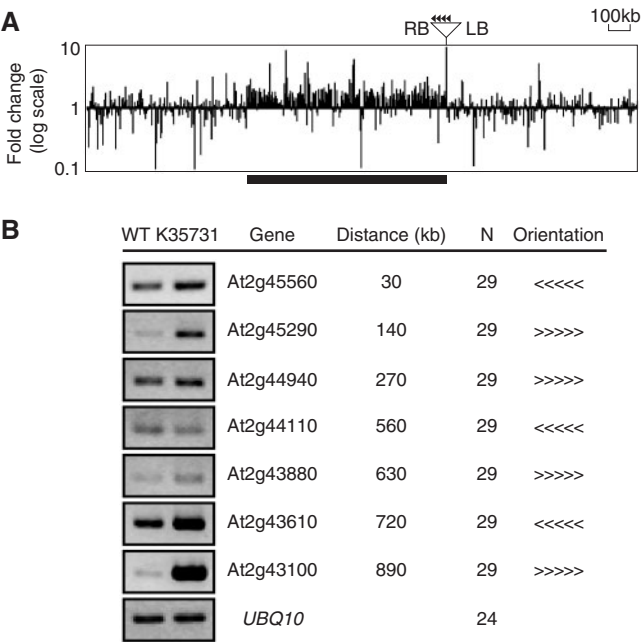


Fig. 4 Genomic response of transcriptional activation by the 35S enhancers in the K35731 mutant. (A) A 2.5 Mb region around a T-DNA insertion locus was shown. A vertical axis indicates fold changes of expression levels of genes on a log scale. A predicted T-DNA insertion in the K35731 mutant is shown as an inverted triangle, on which four copies of the 35S enhancers are depicted as small filled triangles. A 984 kb region on the RB side is underlined. (B) Expression levels of seven randomly selected genes within the 984 kb region that are underlined in (A). Distance from the 35S enhancers of a T-DNA insertion to each gene and the number of PCR cycles are shown. Note that, with the exception of At2g44110, the expression levels of six genes that are located significantly distally from a T-DNA insertion are increased. *UBQ10* was used as an internal control. N, the number of PCR cycles

were isolated in the T₁ generation. Among 111 putative mutants, we determined that female gametophytic mutant phenotypes were stably inherited in the next generation in 22 of the mutants. The T-DNA insertion loci of these 22 mutants were then determined (Table 2). These mutants were classified into three groups on the basis of the number of visible nuclei in the female gametophytes. First, mutants harboring embryos possessing 1-nucleate female gametophytes were assigned to group FA. These mutants were considered to have been arrested in stage FG1 (Drews and Yadegari 2002). For example, in the S25-9 mutant, only one nucleus was visible in 50% of the ovules (Fig. 5A) and subsequent mitosis from the nucleus did not occur, whereas other embryos in the same silique developed into subsequent stages. Two T-DNA insertions were identified in a promoter region of the defensin-like family protein (At5g16453) (Silverstein et al. 2005) and an intergenic region between At2g29120 and At2g29125. Secondly,

mutants that showed 2- to 4-nucleate female gametophytes without further cellularization, or that exhibited abnormal nuclei positions were classified into group FB. These mutants were considered to have been arrested in stages FG2–FG4. One example of this was the S27-11 mutant, in which a T-DNA was inserted into the coding region of a myosin XI-B (At1g04160) (Hashimoto et al. 2005) (Fig. 5B). Thirdly, mutants that showed unfused polar nuclei and, thus, failed to generate a secondary endosperm nucleus (2n), were categorized into group FC. One example of this was the S29-23 mutant, in which a T-DNA insertion had been predicted within the coding region of the translation initiation factor 3 family protein (At1g34360) (Fig. 5C).

Discussion

We have isolated 151 mutants that showed developmental defects from 80,650 activation-tagged mutants. We used TAIL-PCR to identify flanking sequences of T-DNAs and determined T-DNA insertion loci of 5,361 mutants. Although multiple insertions were identified efficiently from a single sequencing run, some of these results were judged to represent false positives. One possible explanation for this is that the incomplete removal of oligonucleotides may induce false positives. An alternative explanation is that the chimeric results arise from template switching, which is induced by *Taq* polymerase (Paabo et al. 1990). This template switching is, however, less likely in our library, as the flanking sequences of a second insertion were recognized from lower amplitude peaks masked by higher amplitude peaks.

It appeared that integration into the plant genome was not completely random, thereby giving rise to regions that contain a large number of T-DNA insertions. These ‘hot spot’ regions in our library were, interestingly, different from those reported previously (Sessions et al. 2002, Szabados et al. 2002, Ichikawa et al. 2003). For example, the q arm ends of chromosomes 1 and 2, which were identified as hot spot regions in the SAIL line (Sessions et al. 2002), did not harbor an unusually high number of insertions in our library. This discrepancy may be attributed to the different levels of euchromatin content in the plants used for infiltration (Wu et al. 2002), or to differences in the favored integration sites of *Agrobacterium* strains, probably depending on differences in the *vir* gene products (Gelvin 2000).

The 35S enhancers may activate multiple genes within a broad range

We have shown that transcription of multiple genes was activated by the 35S enhancers in K35731 mutants (Fig. 4). Two possible explanations exist for this multiple gene activation. First, multiple activation may result from

Table 2 T-DNA insertion loci of female gametophytic mutants identified in this study

Group	Line	No. of insertions	<i>Gene L1</i>	Gene inserted by a T-DNA	<i>Gene R1</i>	Orientation of a T-DNA ^a
FA	S24-18	1	At2g19450		At2g19460	+
	S25-3	1	At1g24470		At1g24480	+
	S25-9	2	At5g16453		At5g16460	+
			At2g29120		At2g29125	—
	S31-13	1	At5g64680		At5g64685	—
	S31-19	2	At3g46750	At3g46740	At3g46730	—
			At5g64680		At5g64685	—
	S31-21	2	At4g23430		At4g23440	—
			At4g38710	At4g38730	At4g38740	+
	S32-2	1	At2g31650		At2g31660	+
	S37-10	1	At1g78760		At1g78750	+
FB	S24-6	1	At1g77910		At1g77920	—
	S27-11	1	At1g04170	At1g04160	At1g04150	+
	S27-22	1	At5g62430		At5g62430	—
	S28-12	1	At2g26700	At2g26710	At2g26720	—
	S28-31	1	At3g26980		At3g27010	—
	S29-4	2	At4g19600	At4g19610	At4g19620	+
			At5g08690		At5g08710	+
	S29-12	1	At2g02380	At2g02390	At2g02400	—
	S29-13	1	At5g55480	At5g55490	At5g55500	+
	S29-22	2	At2g02380	At2g02390	At2g02400	—
			At3g45930		At3g45940	—
	S30-16	1	At2g31230		At2g31240	+
	S30-34	3	At3g13420		At3g13430	+
			At5g66620	At5g66630	At5g66630	+
			At1g35500		At1g35490	+
FC	S31-4	1	At1g21990		At1g22000	+
	S24-15	2	At4g02000	At4g02010	At4g02020	—
			At1g80770		At1g80780	+
	S29-23	1	At1g34370	At1g34360	At1g34355	+

^a A plus (+) indicates that a T-DNA is inserted, such that the 35S enhancers and the RB faced toward the q end of a chromosome. A minus (–) denotes the opposite orientation.

the scanning action of an activator bound to the 35S enhancers, such that an activator tracks the genomic region, inducing transcriptional enhancement of the genes scanned, until it encounters the insulator (Bell and Felsenfeld 1999). Secondly, the activation of multiple genes may be induced by a change in chromatin structure. When inserted into the plant genome, 35S enhancers may induce structural changes around a T-DNA insertion site, thereby releasing the repression of the genes within the region. This notion is consistent with the finding that the enhancer may function, at least in part, to counteract repression induced by chromatin-associated proteins (Kamakaka et al. 1993). Nonetheless, as the transcriptional activation of multiple genes was observed in a single experiment, more data should be accumulated in order to elucidate the mode of action of the 35S enhancers in activation tagging screens.

It should also be noted that the up-regulation of genes by the 35S enhancers in the K35731 mutant was detected only on the RB side. This is interesting, considering that general enhancers are known to act independently of orientation. If the activation of multiple genes can be explained by the scanning hypothesis, the finding that activation occurs only on the RB side bears further examination. Further analysis will also be required to determine precisely the manner in which the 35S enhancers function in activation tagging mutants.

Female gametophytic mutants in our library

Among 22 mutants the mutant phenotypes of which had been inherited in the T₂ generation, T-DNA insertions appeared to disrupt gene structures in 11 mutants. This high rate of T-DNA insertion within the coding region raised the

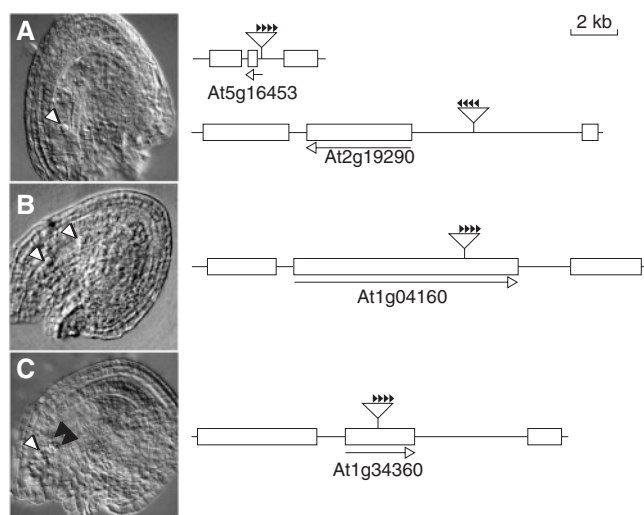


Fig. 5 Phenotypes of female gametophytic mutants isolated in this study and their T-DNA insertion loci. T-DNA insertion maps are shown at the right of each mutant, on which predicted T-DNA insertions are shown as inverted triangles. Four copies of the 35S enhancers are depicted as small filled triangles. Three genes around a T-DNA insertion locus are shown in each map. Open boxes indicate gene structures, and arrows indicate the orientation of a gene whose expression may be affected by a T-DNA insertion. Open arrowheads in the micrographs indicate the nuclei in female gametophytes in each mutant. (A) Fifty per cent of female gametophytes in the S25-9 mutant contained an embryo arrested at the 1-nucleus stage. (B) Fifty per cent of female gametophytes in the S27-11 mutant harbored an embryo arrested in the 2-nucleate stage. (C) Fifty per cent of female gametophytes in the S29-23 mutant contained an embryo, in which polar nuclei failed to fuse. Closed arrowheads in a micrograph indicate two unfused polar nuclei.

possibility that the female gametophytic phenotypes we observed may have been induced by the loss-of-function of a gene. This notion is bolstered by the observation that dominant embryo-lethal mutations cannot be maintained without the involvement of an inducible expression system (Fitzmaurice et al. 2002). Indeed, in one of female gametophyte mutants isolated in this study, S31-19, we found that a T-DNA was inserted in the coding region of At3g46740, whose loss of function was previously identified as embryo-lethal (Baldwin et al. 2005). At3g46740 encodes Toc75-III, which serves as a protein translocation channel at the outer envelope membrane of plastids. It was shown that embryos of homozygous *atToc75-III* knockout mutants aborted at the 2-cell stage, with additional defects in the suspensor and the free nuclear endosperm. Lethality observed in S31-19 and *atToc75-III* mutants suggested that the translocation channel protein plays a critical role in the protein import mechanism in the plant. In addition, these data suggested that our screening strategy using the criteria of the number of visible nuclei

and segregation distortion may prove successful to identify embryo-lethal genes on a large scale, as previously suggested (Christensen et al. 1998).

However, it should be noted that some of the female gametophytic mutants identified in this study may have been false positives. For example, although the group FA mutants were considered to have been arrested in stage FG1, it remains possible that the arrested phenotype at the 1-nucleate stage might have resulted from defects in mitosis, cell cycle regulation or developmental control (Christensen et al. 1998). Mutants showing defective nucleolar division or mutants manifesting abnormal nucleolar fusion can also result in false positives, or can bias our classification system. Finally, mutants showing asynchronous development might alter our identification scheme, considering that female gametogenesis is largely synchronous in wild-type *Arabidopsis* plants (Christensen et al. 1997).

Applications and limitations of this study

In addition to the mutant screening performed in this study, our activation tagging library may prove useful in several contexts. First, our mutant population can be used to screen for conditional mutants that do not, under normal conditions, display a mutant phenotype. Indeed, a conditional mutant that is stress resistant and a mutant that is insensitive to methyl jasmonate were identified in our library (K.H. Paek and Y.D. Choi, personal communication). Secondly, our mutant population can be used for reverse genetic protocols. Since T-DNA insertion loci were mapped and the information of genes adjacent to a T-DNA insertion were integrated into a database, one can search for mutants of interest via examinations of specific genes of interest. Thirdly, the investigation of mutants shown to be defective in female gametophyte development might add to our understanding of seed development, and might also eventually bolster our knowledge of apomixes, an important process related to increased food production (Bicknell and Koltunow 2004).

Although we demonstrated that activation tagging may prove useful in the identification of developmental mutants, it should be noted that activation tagging is not sufficient for the successful identification of certain mutants. This limitation becomes apparent when the overexpression of a gene induced lethality or developmental arrest. For example, an early lethal phenotype was observed in plants constitutively expressing *Arabidopsis* cleavage and polyadenylation specificity factor (*AtCPSF73-I*), which plays a role in mRNA 3'-end processing (Xu et al. 2006). Arrest of growth in the shoot apical meristem was observed in transgenic plants overexpressing *CLAVATA3/ESR* (*CLE*) (Strabala et al. 2006). Therefore, it appears likely that these genes cannot be identified via activation tagging, as no viable mutants can be obtained from the

overexpression of these genes. A possible alternative method that can be used to overcome such problems involves the use of an inducible gene expression system employing dexamethasone or ethanol (Aoyama and Chua 1997, Roslan et al. 2001). In spite of the fact that the overexpression of the gene induced lethality or growth arrest, viable mutants can be maintained unless the genes are activated. Indeed, these conditional expression systems can be used successfully to identify mutants which cannot be readily identified by activation tagging (Deveaux et al. 2003), which suggests that a genetic screening technique using the inducible system may serve to complement activation tagging.

In conclusion, we have applied an activation tagging strategy in this study, in an attempt to isolate mutants that displayed developmental defects. Mutants manifesting dominant developmental abnormalities and female gametophytic mutants were isolated. Our mutant library adds more mutants to the previously generated and currently available mutant reservoirs (Howden et al. 1998, Matzk et al. 2000, Weigel et al. 2000, Marsch-Martinez et al. 2002, Tzafrir et al. 2003, Pagnussat et al. 2005), and may prove useful in the isolation of genes that play roles in plant development.

Materials and Methods

Generation of an activation tagging mutant population

Wild-type *Arabidopsis* plants (ecotype Columbia), grown under long-day conditions (16 h/8 h light/dark) at 23°C, were employed for plant transformation. pSKI015, an activation tagging vector (Weigel et al. 2000), was introduced into *A. tumefaciens* strain GV3101 (pMP90RK) for infiltration. Transgenic plants were selected for herbicide (ammonium glufosinate) resistance. Seeds harvested from individual mutants and seeds pooled from 100 mutants or 1,000 mutants were then deposited in the stock center of the Crop Functional Genomics Center (<http://cfgc.snu.ac.kr/english/index.html>).

Sequence analysis

Sequences flanking T-DNA insertion loci were amplified using a modified TAIL-PCR (Liu and Whittier 1995). Genomic DNA was prepared from the crude extracts using Plant DNeasy 96 DNA Extraction kits (Qiagen, Germany), in accordance with the manufacturer's instructions. TAIL-PCR was conducted using DEG1 primer (5'-WGCNAGTNAGWANAAG-3'); however, for those that did not yield amplified products with DEG1 primer, DEG2 primer (5'-AWGCANGNCWGANATA-3') was employed in order to amplify the flanking sequences. To improve sequencing efficiency, these amplified fragments were sequenced directly without the gel purification of individual bands, after the removal of contaminating oligonucleotides and dNTPs (Nilsen et al. 2001). AtLB4 (5'-GTAGATTTCCTCGGACATGAA-3'), another nested primer located closer to the LB, was employed as a sequencing primer. Sequencing results that did not harbor a T-DNA-*Arabidopsis* genomic DNA junction, or that contained <30 bp of *Arabidopsis* genomic sequences were discarded in an effort to augment accuracy. When nearly identical integration sites were identified from mutants that were processed in a batch, such results

were considered evidence of cross-contamination and were thus discarded. A BLASTi (interactive BLAST) program was developed in order to acquire information of the genes adjacent to the RB and LB of T-DNAs. The positions of the T-DNA insertion sites were mapped with the MapViewer of the TAIR (<http://www.arabidopsis.org/servlets/mapper>).

Expression analysis

Either semi-quantitative RT-PCR or Northern hybridization was employed in order to determine whether the expression of a gene adjacent to a T-DNA insertion site had been up-regulated. Either actin or the *UBQ10* gene was used as a positive control. For a microarray analysis, Affymetrix ATH1 GeneChips were used. Total RNA was extracted using an RNeasy plant kit (Qiagen, Germany), from 10-day-old *Arabidopsis* seedlings grown under long-day conditions at 23°C. The expression data acquired were then analyzed via GeneSpring software (Silicon Genetics, Redwood, CA, USA).

Mutant screen to identify female gametophyte mutants

Maturing ovules in a silique of a plant in the T₁ generation prior to desiccation were manually opened and examined. Two screening methods were utilized in the identification of the female gametophytic mutants. First, putative mutants, in which approximately 50% of ovules were white in the maturing siliques, and thus failed to set seeds, were collected. Secondly, heterozygous female gametophytic mutations were confirmed via determinations of whether they demonstrated an altered segregation ratio for BASTA resistance when self-pollinated. Only those mutants that yielded progeny in a 1:1 ratio (BASTA resistant:BASTA sensitive) were collected. Otherwise, mutants showing 3:1 (BASTA resistant:BASTA sensitive) ratios were considered to be sporophytic mutations and were, therefore, discarded. Mutants were classified in accordance with the numbers of visible nuclei in the female gametophytes (Drews and Yadegari 2002).

Supplementary material

Supplementary material mentioned in the article is available to online subscribers at the journal website www.pcp.oxfordjournals.org.

Acknowledgments

We are grateful to J. Lee and D.-Y. Choe for developing the BLASTi program. We thank Y.J. Kim, O.J. Lee and M.R. Lee for growing the plants. This work was supported by a grant from the Crop Functional Genomics Center to J.S.L.

References

- Aoyama, T. and Chua, N.H. (1997) A glucocorticoid-mediated transcriptional induction system in transgenic plants. *Plant J.* 11: 605–612.
- Arabidopsis Genome Initiative. (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815.
- Baldwin, A., Wardle, A., Patel, R., Dudley, P., Park, S.K., Twell, D., Inoue, K. and Jarvis, P. (2005) A molecular-genetic study of the *Arabidopsis* *Toc75* gene family. *Plant Physiol.* 138: 715–733.
- Bell, A.C. and Felsenfeld, G. (1999) Stopped at the border: boundaries and insulators. *Curr. Opin. Genet. Dev.* 9: 191–198.
- Bicknell, R.A. and Koltunow, A.M. (2004) Understanding apomixis: recent advances and remaining conundrums. *Plant Cell* 16: S228–S245.
- Chalfun-Junior, A., Mes, J.J., Mlynarova, L., Aarts, M.G. and Angenent, G.C. (2003) Low frequency of T-DNA based activation tagging in *Arabidopsis* is correlated with methylation of CaMV 35S enhancer sequences. *FEBS Lett.* 555: 459–463.

- Charlesworth, D. and Charlesworth, B. (1979) The evolutionary genetics of sexual systems in flowering plants. *Proc. R. Soc. B: Biol. Sci.* 205: 513–530.
- Chory, J., Ecker, J.R., Briggs, S., Caboche, M., Coruzzi, G.M., et al. (2000) National Science Foundation-Sponsored Workshop Report: 'The 2010 Project' functional genomics and the virtual plant. A blueprint for understanding how plants are built and how to improve them. *Plant Physiol.* 123: 423–426.
- Christensen, C.A., King, E.J., Jordan, J.R. and Drews, G.N. (1997) Megagametogenesis in *Arabidopsis* wild type and the *Gf* mutant. *Sex. Plant Reprod.* 10: 49–64.
- Christensen, C.A., Subramanian, S. and Drews, G.N. (1998) Identification of gametophytic mutations affecting female gametophyte development in *Arabidopsis*. *Dev. Biol.* 202: 136–151.
- Deveaux, Y., Peaucelle, A., Roberts, G.R., Coen, E., Simon, R., Mizukami, Y., Traas, J., Murray, J.A., Doonan, J.H. and Laufs, P. (2003) The ethanol switch: a tool for tissue-specific gene induction during plant development. *Plant J.* 36: 918–930.
- Drews, G.N. and Yadegari, R. (2002) Development and function of the angiosperm female gametophyte. *Annu. Rev. Genet.* 36: 99–124.
- Fitzmaurice, W.P., Holzberg, S., Lindbo, J.A., Padgett, H.S., Palmer, K.E., Wolfe, G.M. and Pogue, G.P. (2002) Epigenetic modification of plants with systemic RNA viruses. *Omic*s. 6: 137–151.
- Friedrichsen, D.M., Joazeiro, C.A., Li, J., Hunter, T. and Chory, J. (2000) *Brassinosteroid-insensitive-1* is a ubiquitously expressed leucine-rich repeat receptor serine/threonine kinase. *Plant Physiol.* 123: 1247–1256.
- Gelvin, S.B. (2000) *Agrobacterium* and plant genes involved in T-DNA transfer and integration. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 51: 223–256.
- Hashimoto, K., Igarashi, H., Mano, S., Nishimura, M., Shimmen, T. and Yokota, E. (2005) Peroxisomal localization of a myosin XI isoform in *Arabidopsis thaliana*. *Plant Cell Physiol.* 46: 782–789.
- Howden, R., Park, S.K., Moore, J.M., Orme, J., Grossniklaus, U. and Twell, D. (1998) Selection of T-DNA-tagged male and female gametophytic mutants by segregation distortion in *Arabidopsis*. *Genetics* 149: 621–631.
- Ichikawa, T., Nakazawa, M., Kawashima, M., Muto, S., Gohda, K., et al. (2003) Sequence database of 1172 T-DNA insertion sites in *Arabidopsis* activation-tagging lines that showed phenotypes in T1 generation. *Plant J.* 36: 421–429.
- Kamakaka, R.T., Bulger, M. and Kadonaga, J.T. (1993) Potentiation of RNA polymerase II transcription by Gal4-VP16 during but not after DNA replication and chromatin assembly. *Genes Dev.* 7: 1779–1795.
- Kim, J., Jung, J.H., Reyes, J.L., Kim, Y.S., Kim, S.Y., et al. (2005) microRNA-directed cleavage of *ATHB15* mRNA regulates vascular development in *Arabidopsis* inflorescence stems. *Plant J.* 42: 84–94.
- Liljegren, S.J., Ditta, G.S., Eshed, Y., Savidge, B., Bowman, J.L. and Yanofsky, M.F. (2000) *SHATTERPROOF* MADS-box genes control seed dispersal in *Arabidopsis*. *Nature* 404: 766–770.
- 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.
- Matzk, F., Meister, A. and Schubert, I. (2000) An efficient screen for reproductive pathways using mature seeds of monocots and dicots. *Plant J.* 21: 97–108.
- Moss, E.G. (2002) MicroRNAs: hidden in the genome. *Curr. Biol.* 12: R138–R140.
- Nakazawa, M., Ichikawa, T., Ishikawa, A., Kobayashi, H., Tshara, Y., Kawashima, M., Suzuki, K., Muto, S. and Matsui, M. (2003) Activation tagging, a novel tool to dissect the functions of a gene family. *Plant J.* 34: 741–750.
- Naoki, K. and Hashimoto, T. (2004) A semidominant mutation in an *Arabidopsis* mitogen-activated protein kinase phosphatase-like gene compromises cortical microtubule organization. *Plant Cell* 16: 1841–1853.
- Nilsen, I.W., Overbo, K. and Olsen, R.L. (2001) Thermolabile alkaline phosphatase from Northern shrimp (*Pandalus borealis*): protein and cDNA sequence analyses. *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* 129: 853–861.
- Ostergaard, L. and Yanofsky, M.F. (2004) Establishing gene function by mutagenesis in *Arabidopsis thaliana*. *Plant J.* 39: 682–696.
- Paabo, S., Irwin, D.M. and Wilson, A.C. (1990) DNA damage promotes jumping between templates during enzymatic amplification. *J. Biol. Chem.* 265: 4718–4721.
- Pagnussat, G.C., Yu, H.J., Ngo, Q.A., Rajani, S., Mayalagu, S., Johnson, C.S., Capron, A., Xie, L.F., Ye, D. and Sundaresan, V. (2005) Genetic and molecular identification of genes required for female gametophyte development and function in *Arabidopsis*. *Development* 132: 603–614.
- Palatnik, J.F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J.C. and Weigel, D. (2003) Control of leaf morphogenesis by microRNAs. *Nature* 425: 257–263.
- Perrella, G., Cremona, G., Consiglio, F., Errico, A., Bressan, R.A. and Conicella, C. (2006) Screening for mutations affecting sexual reproduction after activation tagging in *Arabidopsis thaliana*. *J. Appl. Genet.* 47: 109–111.
- Pierik, R., Whitelam, G.C., Voeseek, L.A., de Kroon, H. and Visser, E.J. (2004) Canopy studies on ethylene-insensitive tobacco identify ethylene as a novel element in blue light and plant–plant signalling. *Plant J.* 38: 310–319.
- Roslan, H.A., Salter, M.G., Wood, C.D., White, M.R., Croft, K.P., et al. (2001) Characterization of the ethanol-inducible alc gene-expression system in *Arabidopsis thaliana*. *Plant J.* 28: 225–235.
- Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., et al. (2004) The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Res.* 32: 5539–5545.
- Schneider, A., Kirch, T., Gigolashvili, T., Mock, H.P., Sonnwald, U., Simon, R., Flugge, U.I. and Werr, W. (2005) A transposon-based activation-tagging population in *Arabidopsis thaliana* (TAMARA) and its application in the identification of dominant developmental and metabolic mutations. *FEBS Lett.* 579: 4622–4628.
- Seki, H., Nishizawa, T., Tanaka, N., Niwa, Y., Yoshida, S. and Muranaka, T. (2005) Hairy root-activation tagging: a high-throughput system for activation tagging in transformed hairy roots. *Plant Mol. Biol.* 59: 793–807.
- Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., et al. (2002) A high-throughput *Arabidopsis* reverse genetics system. *Plant Cell* 14: 2985–2994.
- Silverstein, K.A., Graham, M.A., Paape, T.D. and VandenBosch, K.A. (2005) Genome organization of more than 300 defensin-like genes in *Arabidopsis*. *Plant Physiol.* 138: 600–610.
- Strabala, T.J., O'Donnell, P. J., Smit, A.M., Ampomah-Dwamena, C., Martin, E.J., Netzler, N., Nieuwenhuizen, N.J., Quinn, B.D., Foote, H.C. and Hudson, K.R. (2006) Gain-of-function phenotypes of many *CLAVATA3/ESR* genes, including four new family members, correlate with tandem variations in the conserved *CLAVATA3/ESR* domain. *Plant Physiol.* 140: 1331–1344.
- Szabados, L., Kovacs, I., Oberschall, A., Abraham, E., Kerekcs, I., et al. (2002) Distribution of 1000 sequenced T-DNA tags in the *Arabidopsis* genome. *Plant J.* 32: 233–242.
- Tani, H., Chen, X., Nurnberg, P., Grant, J.J., SantaMaria, M., Chini, A., Gilroy, E., Birch, P.R. and Loake, G.J. (2004) Activation tagging in plants: a tool for gene discovery. *Funct. Integr. Genomics* 4: 258–266.
- Tzafrir, I., Dickerman, A., Brazhnik, O., Nguyen, Q., McElver, J., Frye, C., Patton, D. and Meinke, D. (2003) The *Arabidopsis* SeedGenes Project. *Nucleic Acids Res.* 31: 90–93.
- Vizir, I.Y., Anderson, M.L., Wilson, Z.A. and Mulligan, B.J. (1994) Isolation of deficiencies in the *Arabidopsis* genome by gamma-irradiation of pollen. *Genetics* 137: 1111–1119.
- Weigel, D., Ahn, J.H., Blazquez, M.A., Borevitz, J.O., Christensen, S.K., et al. (2000) Activation tagging in *Arabidopsis*. *Plant Physiol.* 122: 1003–1013.
- Wu, J., Maehara, T., Shimokawa, T., Yamamoto, S., Harada, C., et al. (2002) A comprehensive rice transcript map containing 6591 expressed sequence tag sites. *Plant Cell* 14: 525–535.
- Xu, R., Zhao, H., Dinkins, R.D., Cheng, X., Carberry, G. and Li, Q.Q. (2006) The 73 kD subunit of the cleavage and polyadenylation specificity factor (CPSF) complex affects reproductive development in *Arabidopsis*. *Plant Mol. Biol.* 61: 799–815.
- Yadegari, R. and Drews, G.N. (2004) Female gametophyte development. *Plant Cell* 16: S133–S141.
- Zhao, Y., Christensen, S.K., Fankhauser, C., Cashman, J.R., Cohen, J.D., Weigel, D. and Chory, J. (2001) A role for flavin monooxygenase-like enzymes in auxin biosynthesis. *Science* 291: 306–309.

(Received August 14, 2006; Accepted November 29, 2006)