FEBS 27882

FEBS Letters 555 (2003) 459–463

Low frequency of T-DNA based activation tagging in *Arabidopsis* is correlated with methylation of CaMV 35S enhancer sequences

Antonio Chalfun-Junior[a], Jurriaan J. Mes[a], Ludmila Mlynárová[a], Mark G.M. Aarts[b], Gerco C. Angenent[a,*]

[a] *Business Unit Bioscience, Plant Research International, Bornsesteeg 65, 6708 PD Wageningen, The Netherlands*
[b] *Laboratory of Genetics, Wageningen University, Arboretumlaan 4, 6703 BD Wageningen, The Netherlands*

Received 22 May 2003; revised 25 August 2003; accepted 30 October 2003

First published online 18 November 2003

Edited by Marc Van Montagu

Abstract **A powerful system to create gain-of-function mutants in plants is activation tagging using T-DNA based vehicles to introduce transcriptional enhancer sequences. Large *Arabidopsis* populations of individual plants carrying a quadruple cauliflower mosaic virus (CaMV) 35S enhancer are frequently used for mutant screenings, however the frequency of morphological mutants remains very low. To clarify this low frequency we analyzed a subset of lines generated by this method. The correlation between the number of T-DNA insertion sites, the methylation status of the 35S enhancer sequence and 35S enhancer activity was determined. All plants containing more than a single T-DNA insertion showed methylation of the 35S enhancer and revealed a dramatic decrease in 35S enhancer activity. The results support the notion that in a large proportion of the T-DNA based activation tagged lines the 35S transcriptional enhancer is silenced due to methylation, which is induced by multiple T-DNA integrations.**

Key words: Activation tagging; *Arabidopsis* mutant; Transcriptional gene silencing; DNA methylation

1. Introduction

At present many approaches are available to create genetic variants, like mutants, that may give clues about plant processes and their genetic control. Loss-of-function mutants are the most informative, because the mutant phenotype is directly associated with the biological function of the impaired gene. However, the use of loss-of-function mutants has its limitations, e.g. in the case of gene redundancy. This is a general phenomenon caused by genome duplications as is clearly present in *Arabidopsis* [1], where the mutation of one gene copy often does not result in a detectable phenotype when other copies are still functional. An alternative approach for gene function analysis is gain-of-function mutagenesis, in which a gene is either ectopically or constitutively overexpressed compared to normal expression levels or patterns [2]. Although in such a case a particular gene is not expressed in its normal biological context, it may give clues about how a gene affects certain biological processes. Ectopic expressions and overexpressions of individual genes have been performed for more than a decade and have provided valuable information about gene functions [3]. An alternative to the gene-by-gene overexpression approach was developed by Walden et al. [4], who designed a T-DNA based activation tagging approach to identify and isolate novel genes from tobacco. Only recently large-scale activation tagging gain-of-function mutagenesis was reported for *Arabidopsis* [5]. Weigel and co-workers [5] described the establishment of a large collection of *Arabidopsis* plants, transformed with a binary vector carrying a quadruple version of the cauliflower mosaic virus (CaMV) 35S enhancer. This T-DNA activation tagging system has since then been used by many researchers and the obtained results illustrate the strength of the technology [6–8].

Although very successful, the frequency of dominant morphological mutants was surprisingly only around 1 in 1000 plants, which is much lower than can be expected considering the densely packed *Arabidopsis* genome and the estimated distance of up to 3 kb between an activated gene and the 4×35S enhancer [5].

This relatively low frequency of T-DNA activation mutants is in contrast to recently published results obtained with a transposon based activation tagging system [9]. The latter system, which generates single insertions of the activation tag, yields a frequency of about 1% morphological mutants, thus 10-fold higher than obtained with T-DNA based activation tagging [9].

Despite the low activation tagging frequency, the T-DNA based activation tagging approach is highly popular in *Arabidopsis* mutant screens [10,11] with new populations being generated [12]. Nevertheless, with a low activation frequency, the method appears to be less attractive for a number of applications, for instance the use in species which cannot be easily transformed in similarly large numbers as *Arabidopsis*. It will also be less suitable for more complicated or laborious *Arabidopsis* screens or for screens where only a few mutants are expected due to the specificity of the mutant class. We were therefore curious to discover the reason behind the low activation tagging frequency. The results described in this paper indicate that enhancer methylation and subsequent transcriptional silencing is a plausible explanation for the relatively low mutation frequency in T-DNA based activation tagging populations.

*Corresponding author. Fax: (31)-317-423110.
E-mail address: gerco.angenent@wur.nl (G.C. Angenent).

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)01300-0

2. Materials and methods

2.1. Plant material

The four seed batches originated from set 2 of the 'Weigel' collection, ecotype Columbia (Col-7) and were obtained from the Nottingham *Arabidopsis* Stock Centre, Nottingham, UK (http://nasc.nott.ac.uk/). Plants were grown in soil, under normal greenhouse conditions (22°C, 14/10 h light/dark). For selection on plates, seeds were surface sterilized by vapor phase seed sterilization (http://plantpath.wisc.edu/~afb/vapster.html) and selection with 10 mg l^{-1} hygromycin and 15 mg l^{-1} phosphinothricin-DL (PTT) was applied on medium.

2.2. Southern blots

Genomic DNA was isolated from rosette leaves and approximately 300 ng of DNA was digested with restriction enzyme *Hin*dIII for the determination of number of T-DNA integrations. For methylation determination, genomic DNA was digested with the isoschizomers *Sau*3AI and *Mbo*I. DNA was electrophoresed in a 1.0% (w/v) agarose gel in 1×TBE (1.0 M Tris, 0.9 M boric acid, 0.01 M ethylenediamine tetraacetic acid (EDTA)) blotted onto Hybond N+ membrane (Amersham Pharmacia Biotech) following the normal instructions of the manufacturer. As probe, a fragment containing the 35S enhancers was used and labeled by random oligonucleotide priming (Gibco BRL®).

2.3. Inverted repeat determination

Inverted repeats were determined by polymerase chain reaction (PCR), using a single primer designed 59 bp upstream of the 4×35S enhancers (PRO015), followed by a nested PCR with a primer overlapping the junction between the copies of the 35S enhancer (PRO016). Visualization was performed by hybridization of the PCR product with a probe containing 35S enhancer sequences. The following oligonucleotide sequences were used: PRO015: 5′-CGACTCACTATAGGGCGAATTGG-3′ and PRO016: 5′-ATGTGATATCTAGATCCCCAACATGG-3′.

2.4. GUS activity determination

For GUS assay, protein extracts were prepared by grinding four just opened flowers, for each replication, in 50 ul extraction buffer (100 mM sodium phosphate pH 6.7, 10 mM EDTA, 0.1% Triton X-100, 0.1% sarcosyl, 10 mM dithiothreitol (DTT)), cleared by centrifugation for 5 min at 4°C, when MUG assay measurements were performed following previous description [13].

2.5. Copy number level

The number of T-DNA integrations was determined by measuring the intensity of all hybridizing bands from a particular plant in the F1 Southern blot. These data were adjusted for unequal loading using the two internal standards, i.e. the bands representing the 35S::GUS construct as a reference. The estimated number of T-DNA integrations were categorized into three groups of different copy number levels: from low, medium to high copy number levels. The F1 Southern blot was scanned with a Bio-Rad Molecular Imager® FX, and quantification analysis was done by Quantity One® software, version 4, Bio-Rad.

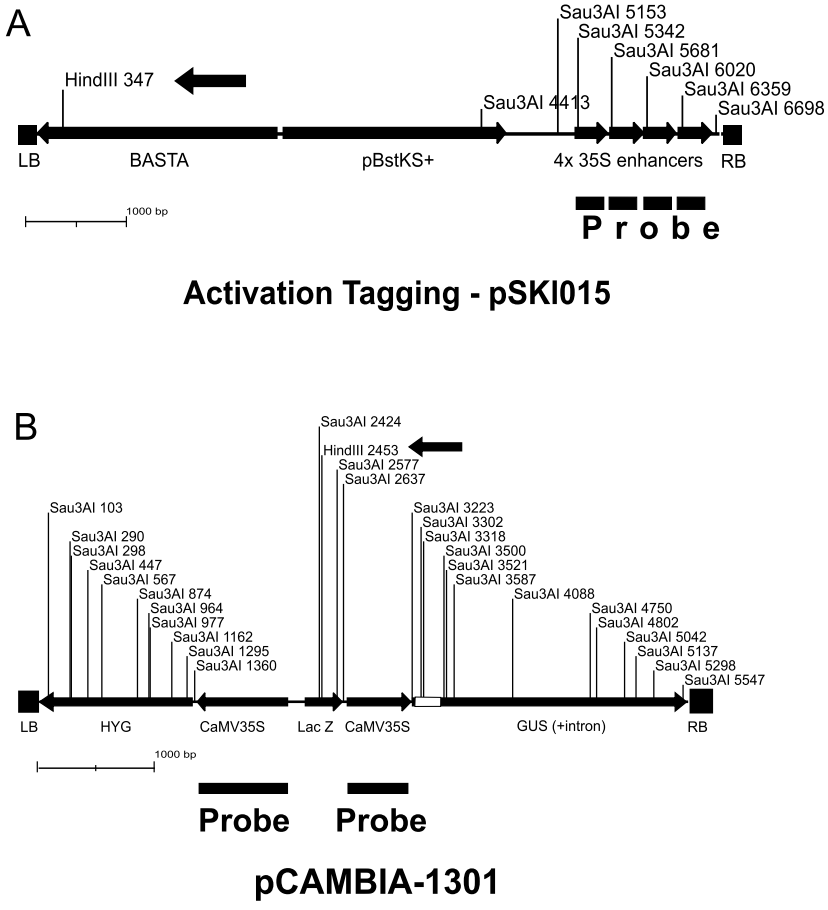


Fig. 1. Schematic representation of the two T-DNA constructs used. Numbers shown with the restriction enzyme sites are the position where the respective enzyme cuts, counting from the left border (LB) towards right border (RB). A: Activation tagging (pSKI015) T-DNA. To facilitate visualization, only the *Sau*3AI/*Mbo*I sites present in the 35S enhancer region are depicted. The unique *Hin*dIII cleavage site is shown and highlighted by a black arrow. The four black square blocks are the regions used as probe, representing sequences of the 35S enhancer. B: pCAMBIA1301 T-DNA. All *Sau*3AI/*Mbo*I sites are shown. The unique *Hin*dIII cleavage site is shown and indicated by a black arrow. White box in the GUS gene represents the catalase intron. Two fat lines are the sequences used as probe (35S sequences).

3. Results and discussion

3.1. Many T-DNA copies are present in activation tagging lines

We analyzed a small randomly chosen subset of plants from the T-DNA activation tagging population obtained from the *Arabidopsis* Stock Centre [5]. From four different seed batches (from pools of 96 plants), 12 T4 plants per seed batch were sown and labeled as follows: plants 1–12 from seed batch N23142, 13–24, from N23077, 25–36 from N23110 and 37–48 from N23190. A phenotypic analysis revealed one plant (#28) with an aberrant phenotype. This plant had a sterile, bushy and dwarfed phenotype and was unfortunately lost before enough tissue was collected for DNA isolation.

Southern blot analysis was performed for the remaining 47 independent T4 plants and a Columbia (Col-0) wild-type plant to get an idea about the number of T-DNA insertions, using *Hin*dIII as a restriction enzyme and 35S fragment sequences as a probe. Among the 47 plants tested, 33 had a T-DNA integrated, with 26 apparently having multiple insertion sites and seven with single or low number of insertions (data not shown). From the 33 plants, four plants with many insertions (plants #10, 11, 20 and 38) and seven with low copy numbers (plants #2, 5, 15, 16, 21, 40 and 46) were chosen for further experiments.

Based on previous reports [14–17] there may be a direct correlation between the number of T-DNA insertions and the methylation status of the T-DNA inserts, in particular when these are present as inverted repeats.

Other studies suggest that the complex T-DNA integration pattern often generated during plant transformations may trigger silencing mechanisms [18,19]. In particular inverted repeats of T-DNA copies can induce either transcriptional gene silencing, through methylation of promoter sequences [20,21], or posttranscriptional gene silencing causing a selective break-down of mRNA molecules [22,23].

We hypothesized that methylation is the reason for the relatively low activation capacity of the 35S enhancers used in the T-DNA activation tagged lines. To obtain evidence for this hypothesis the methylation status of the quadruple 35S enhancer was determined. Southern blot analysis was performed with DNA digested with methylation sensitive and non-sensitive restriction enzymes *Sau*3AI and *Mbo*I, respectively, and with 35S enhancer sequences as a probe. Except for plant #16, all plants showed methylation of the 35S enhancer (data not shown).

A T5 offspring was raised from the 11 selected T4 plants for further analysis of the methylation and for the analysis of CaMV 35S promoter activity. One T-DNA containing T5 plant derived from each T4 parent and a Col wild-type plant were crossed with a plant containing a single copy of the pCAMBIA1301 T-DNA carrying a 35S::GUS construct [24]. F1 offspring plants were selected for the presence of both the activation tag T-DNA and the 35S::GUS constructs (Fig. 1A,B). One F1 plant per cross was analyzed again by Southern blot to estimate the number of T-DNA insertions (Fig. 2). At the same time, we determined the presence of inverted sequences of the T-DNA right border where the quadruple 35S enhancer is located, in a PCR using one outward directed primer (Fig. 3A). This analysis confirmed the previously estimated number of insertions except for plant #2, which now has a single integration locus of the T-DNA probably due to segregation of the T-DNA loci.

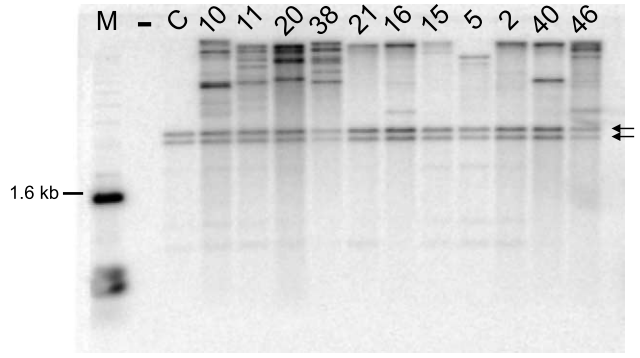


Fig. 2. Southern blot analysis of the 12 F1 plants generated from the cross between 11 selected activation tagging lines (T5) and a plant homozygous for the 35S::GUS construct. Genomic DNA was cut with *Hin*dIII and the blot was probed with 35S sequences. The two fragments indicated with arrows are derived from the 35S::GUS construct. C, control plant (F1 of Col-0×35S::GUS). M, marker lane.

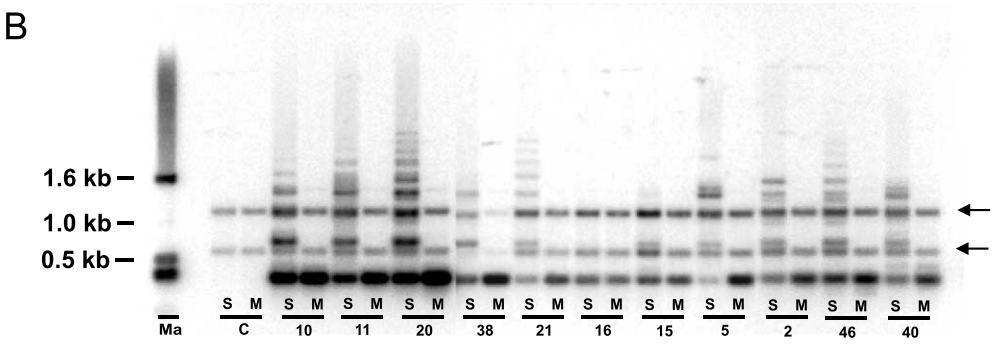
3.2. Methylation of the 4×35S enhancer silences a 35S promoter in trans

Methylation of the quadruple 35S enhancer was determined by Southern blot analysis of the selected F1 plants using methylation sensitive and non-sensitive restriction enzymes *Sau*3AI and *Mbo*I again and 35S enhancer sequences as probe (Fig. 3B). The fragments shared by all F1 plants, including the F1 with wild-type Columbia, represent the 35S promoter present in the pCAMBIA1301 35S::GUS reporter construct. The *Sau*3AI/*Mbo*I sites from these fragments, which are used as internal controls, are located outside the 35S promoter sequences and therefore are not likely to be susceptible to methylation (Fig. 1B). The methylation analysis reveals that all plants except plant #16, which has a single 'activation' T-DNA insertion, show detectable methylation of the 35S enhancer, similar to previously determined for the T4 and T5 generations (data not shown). This confirms our idea that the occurrence of multicopy T-DNA insertions frequently leads to methylation of the 35S enhancer. Based on PCR amplification of the inserts (results not shown) we found that plants #11, 20, 38, and 40 contain an inverted repeat of the T-DNA. These plants show all high levels of methylation, suggesting a correlation between integration structure and methylation. In contrast however, a simple integration structure, e.g. in plant #2 does not always abolish methylation. It must be noted though that the F1 offspring of plant #2 originated from a T4 plant containing at least two T-DNA loci. An explanation therefore might be that the observed methylation is epigenetically inherited from the parental line [25–27]. Alternatively, it may be that occasionally a single integrated copy of the T-DNA is sufficient to trigger methylation of the enhancer, depending on the site of insertion [28–30].

After that we showed frequent occurrence of methylation of the quadruple 35S enhancer we were interested to see if this also had an effect on the activity of the enhancer, as was expected from transcriptional gene silencing experiments [20,21]. This was tested indirectly by measuring GUS activity in the same F1 plants used to estimate copy number level and methylation status (Fig. 3C). All plants, except plant #16, show a significant reduction of GUS expression when compared to the control plant (C; Col×35S::GUS). There appears to be a negative correlation between methylation status

A

PLANT	C	10	11	20	38	21	16	15	5	2	46	40
IR	-	-	+	+	+	-	-	-	-	-	-	+
CNL	-	M	M	H	H	L	L	L	L	L	M	L



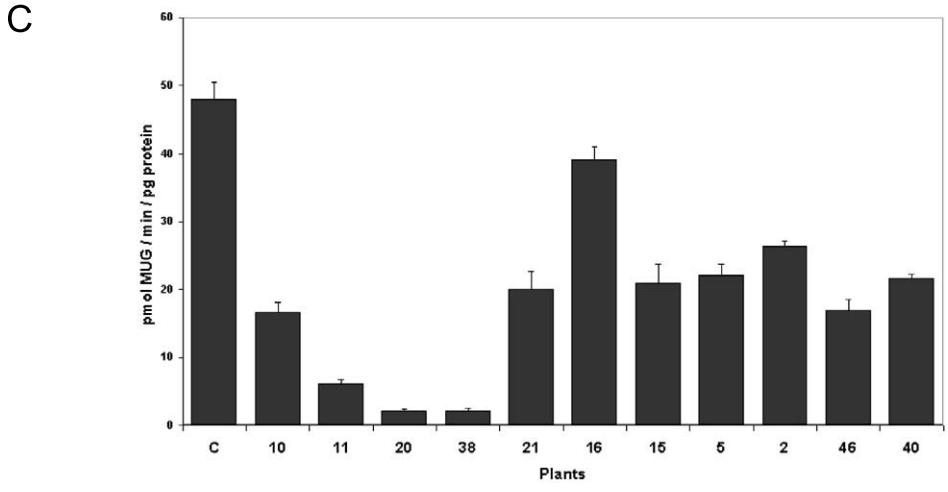


Fig. 3. Analyses of 12 F1 plants generated from the cross of 11 T5 activated tagged plants plus a wild-type plant (Col-0) with a homozygous plant containing a single 35S::GUS insertion. F1 plants obtained from the cross between Col-0 and 35S::GUS are used as control (C). T-DNA copy number levels, methylation status of the 35S enhancer and 35S promoter activity were determined. A: The copy number level (CNL) of the T-DNA activation tagging vector as determined by Southern blot hybridization analyses. The number of T-DNA insertions was determined by measuring the intensity of the hybridizing bands derived from the activation tagging construct (see Fig. 2). The bands derived from the 35S::GUS construct from the pCAMBIA1301 were used as internal control. The plants were categorized in three different 'copy number levels' > five copies). Presence (+) or absence (−) of inverted repeats (IR) determined by PCR using a single primer designed 59 bp upstream of the 4×35S enhancers are indicated. B: Methylation analysis of the F1 plants derived from a cross between activation tagged plants and a homozygous plant with a single copy T-DNA insert containing a 35S::GUS construct (pCAMBIA1301). DNA was digested with *Sau*3AI (S) and *Mbo*I (M), a methylation sensitive and non-sensitive restriction enzyme, respectively. The two bands seen in control (C) represent the pCAMBIA1301 T-DNA insert, and can be used as a loading control. Ma, marker. C: Expression of the GUS transgene in the F1 plants of crosses between 11 different activation tagging T5 plants and Col-0 with a 35S::GUS plant. GUS expression was measured by a MUG assay [13] on four just opened flowers. GUS activity levels were expressed as pmol of methyl-umbelliferone per min per μg of soluble protein (pmol MUG per min per pmol protein). Assays were repeated twice for each plant and the mean values are indicated as bars. The standard deviation of the mean is shown as thin line. C, control plant (F1 of Col-0×35S::GUS).

and GUS activity. This result means that in the presence of methylated copies of the 35S enhancer, a 35S promoter (containing sequences identical to the methylated 35S enhancer sequences) is silenced in trans. Probably the same transcriptional gene silencing mechanism, which caused methylation of the 35S enhancer sequences, causes methylation of the 35S promoter driving GUS expression [31]. Although we did not directly measure silencing of the 35S enhancer, we postulate that the observed methylation is sufficient to reduce the transcription enhancing potential of the quadruple 35S enhancer.

4. Discussion

Silencing of the 35S enhancer as a result of methylation is very likely the reason for the relatively low frequency of mutants found by T-DNA based activation tagging in the 'Weigel' population [5]. We have shown in a representative selec-

tion of lines from this collection that the quadruple 35S enhancer sequence is frequently methylated. The presence of the methylated 35S enhancer sequences resulted in silencing of an in trans copy of a 35S promoter, normally driving transcription of a GUS marker gene.

Several studies report that the presence of inverted repeats in complex chromosomal structures is the main reason for triggering methylation and subsequent silencing of the introduced sequences [16,20,30,32]. We observed relative high levels of methylation in plants with high numbers of T-DNA integrations especially when right border inverted repeats are present. Still, we did find plants in which no inverted repeat was detected, but in which the enhancer was methylated. For these plants other triggers of the methylation machinery could be active, such as the presence of direct repeats, DNA–DNA pairing of the enhancer region or the perception of repeated DNA as foreign [33].

Based on our observations, the T-DNA based activation tagging system might be improved by using a silencing impaired background. The methylation defective *ddm* mutants [26,34] are candidates for such an approach. Recently a number of other silencing defective mutants have been reported which could also be used [35]. As the regulation of transcriptional or posttranscriptional silencing is still not fully understood, probably a number of these mutants should be tried.

Another possibility is the use of single activation tag integrations. This strategy is followed in transposon based activation tagging [9], which has yielded plants with only a single copy of the activating construct. The relative high frequency of morphological mutants (approximately 1%) observed in the transposon based activation tagging population is in line with this single copy hypothesis.

A further improvement of the activation tagging system could be obtained by using an endogenous enhancer, which might be less susceptible to the silencing machinery than a foreign sequence such as the viral 35S enhancer. This would also enable the use of tissue specific enhancers for the ectopic activation of gene expression in a tissue specific manner. The use of such enhancers has not been reported yet, but they can be attractive modifications to the T-DNA/35S enhancer based activation tagging as researchers are using it nowadays.

Acknowledgements: We thank Jan Peter Nap for his help with the MUG assay and useful suggestions and discussions. We also thank Maarten Koornneef for carefully reading the manuscript and stimulating discussions. This work has been financially supported by CAPES-BRAZIL (1519/98-0).

References

[1] The *Arabidopsis* Genome Initiative (2000) Nature 408, 796–815.
[2] Walden, R. and Schell, J. (1994) Agro Food Ind. Hi Tech 5, 9–12.
[3] Hayashi, H., Czaja, I., Lubenow, H., Schell, J. and Walden, R. (1992) Science 258, 1350–1353.
[4] Walden, R., Fritze, K., Hayashi, H., Miklashevischs, E., Harling, H. and Schell, J. (1994) Plant Mol. Biol. 26, 1521–1528.
[5] Weigel, D., Ahn, J.H., Blazquez, M.A., Borevitz, J.O., Christensen, S.K., Fankhauser, C., Ferrandiz, C., Kardailsky, I., Malancharuvil, E.J. and Neff, M.M. et al. (2000) Plant Physiol. 122, 1003–1013.
[6] Borevitz, J.O., Xia, Y., Blount, J., Dixon, R.A. and Lamb, C. (2000) Plant Cell 12, 2383–2393.
[7] Lee, H., Suh, S.S., Park, E., Cho, E., Ahn, J.H., Kim, S.G., Lee, J.S., Kwon, Y.M. and Lee, I. (2000) Genes Dev. 14, 2366–2376.
[8] Ito, T. and Meyerowitz, E.M. (2000) Plant Cell 12, 1541–1550.
[9] Marsch-Martínez, N., Greco, R., van Arkel, G., Herrera-Estrella, L. and Pereira, A. (2002) Plant Physiol. 129, 1544–1556.
[10] van der Graaff, E., Dulk-Ras, A.D., Hooykaas, P.J.J. and Keller, B. (2000) Development 127, 4971–4980.
[11] Huang, S., Cerny, E., Bhat, D.S. and Brown, S.M. (2001) Plant Physiol. 125, 573–584.
[12] Nakazawa, M., Ichikawa, T., Ichikawa, A., Kobayashi, H., Tsuhara, Y., Kawashima, M., Muto, S. and Matsui, M. (2002) in: 13th International Conference on *Arabidopsis* Research, Seville, Spain, abstract, 1/19.
[13] Mlynárová, L., Loonen, A., Heldens, J., Jansen, R.C., Keizer, P., Stiekema, W.J. and Nap, J.P. (1994) Plant Cell 6, 417–426.
[14] Vaucheret, H. and Fagard, M. (2001) Trends Genet. 17, 29–35.
[15] Waterhouse, P.M., Wang, M.B. and Finnegan, E.J. (2001) Trends Plant Sci. 6, 297–301.
[16] Muskens, M.W.M., Vissers, A.P.A., Mol, J.N.M. and Kooter, J.M. (2000) Plant Mol. Biol. 43, 243–260.
[17] Matzke, M.A., Matzke, A.J.M., Pruss, G.J. and Vance, V.B. (2001) Curr. Opin. Gen. Dev. 11, 221–227.
[18] Nacry, P., Camilleri, C., Courtial, B., Caboche, M. and Bouchez, D. (1998) Genetics 149, 641–650.
[19] Tax, F.E. and Vernon, D.M. (2001) Plant Physiol. 126, 1527–1538.
[20] Mette, M.F., Aufsatz, W., van der Winden, J., Matzke, M.A. and Matzke, A.J.M. (2000) EMBO J. 19, 5194–5201.
[21] Sijen, T., Vijn, I., Rebocho, A., van Blokland, R., Roelofs, D., Mol, J.N.M. and Kooter, J.M. (2001) Curr. Biol. 11, 436–440.
[22] Kooter, J.M., Matzke, M.A. and Meyer, P. (1999) Trends Plant Sci. 4, 340–347.
[23] Fagard, M. and Vaucheret, H. (2000) Annu. Rev. Plant Physiol. Plant Mol. Bio. 51, 167–194.
[24] Roberts, C.S., Rajagopal, S., Smith, L.A., Nguyen, T.A., Yang, W., Nugroho, S., Ravi, K.S., Vijayachandra, K., Harcourt, R.L., Dransfield, L., Desamero, N., Slamet, I., Hajdukiewicz, P., Svab, Z., Maliga, P., Mayer, J.E., Keese, P., Kilian, K. and Jefferson, R.A. (2002) http://www.cambia.org.au/main/r_et_vman.htm.
[25] Mittelsten Scheid, O., Afsar, K. and Paszkowski, J. (1998) Proc. Natl. Acad. Sci. USA 95, 632–637.
[26] Kakutani, T., Munakata, K., Richards, E.J. and Hirochika, H. (1999) Genetics 151, 831–838.
[27] Chareonpornwattana, S., Thara, K.V., Wang, L., Datta, S.K., Panbangred, W. and Muthukrishnan, S. (1999) Theor. Appl. Genet. 98, 371–378.
[28] Stam, M., de Bruin, R., van Blokland, R., van der Hoorn, R.A.L., Mol, J.N.M. and Kooter, J.M. (2000) Plant J. 21, 27–42.
[29] Montgomery, M.K. and Fire, A. (1998) Trends Genet. 14, 255–258.
[30] Stam, M., Ada, V., Mol, J.N.M. and Kooter, J.M. (1998) Mol. Cell. Biol. 18, 6155–6177.
[31] Mette, M.F., van der Winden, J., Matzke, M.A. and Matzke, A.J.M. (1999) EMBO J. 18, 241–248.
[32] de Buck, S. and Depicker, A. (2001) Mol. Gen. Genet. 265, 1060–1068.
[33] Matzke, M.A. and Matzke, A.J.M. (1995) Plant Physiol. 107, 679–685.
[34] Martienssen, R.A. and Colot, V. (2001) Science 293, 1070–1074.
[35] Mittelsten Scheid, O. and Paszkowski, J. (2000) Plant Mol. Biol. 43, 235–241.