The 13th Symposium on Plant Biotechnology **13:** 11-15, 1999.
1999 Korea-Japan Join Symposium.

The transposition pattern of the *Ac* element and its use for targeted transposition in *Arabidopsis thaliana*

Yasunori Machida*, Hitoshi Onouchi, Hirokazu Tanaka, Susumu Hamada, Takaaki Ishikawa, Endang Semiarti, Hidekazu Iwakawa, Kiyohito Nomura, and Chiyoko Machida[1]
*Laboratory of Developmental Biology, Division of Biological Science, Graduate School of Science, Nagoya University, Chikusa-ku, Nagoya 464-8602; [1]On leave from Kirin Brewery Co., Ltd. Central Laboratories for Key Technology, 1-13-5 Fukuura, Kanazawa-ku, Yokohama, Kanagawa 236, Japan; *Laboratory of Developmental Biology, Division of Biological Science, Graduate School of Science, Nagoya University, Chikusa-ku, Nagoya 464-01, Japan (author for correspondence)*

Abstract

In order to evaluate feasibility of the gene tagging by the maize transposable element *Ac* in heterologous plant systems, we have investigated physical distances and directions of transposition of the element in *Arabidopsis thaliana* and tobacco cultured cell line BY-*2*. We prepared a T-DNA construct that carried a non-autonomous derivative of *Ac* with a site for cleavage by endonuclease I-*Sce*I (designated dAc-I-RS element). Another cleavage site was also introduced into the T-DNA region outside dAc-I-RS. A number of transgenic Arabidopsis plants were generated, each of which had a single copy of the T-DNA at a different chromosomal location. To examine the pattern of transposition, three out of these transgenic plants were crossed with the Arabidopsis plant that carried the gene for *Ac* transposase and progeny in which dAc-I-RS had been transposed were isolated. After digestion of the genomic DNA of these progeny with I-*Sce*I, sizes of segment of DNA were determined by pulse-field gel electrophoresis. We also performed linkage analysis for the transposed elements and sites of mutations near the elements. Our results with three transgenic lines showed that 50% of all transposition events had occurred within 1,700 kilo-base pairs (kb) on the same chromosome, with 35% within 200 kb, and that the elements transposed in both directions on the chromosome with roughly equal probability. The data thus indicate that the *Ac-Ds* system is most useful for tagging of genes that are present within 200 kb of the chromosomal site of *Ac in Arabidopsis*. In addition, determination of the precise localization of the transposed dAc-I-RS element should definitely assist in map-based cloning of genes around insertion sites. In the present paper, we report typical examples of such gene isolation studies.

Abbreviations: P35S, cauliflower mosaic virus 35S promoter; PFGE, pulse-field gel electrophoresis; HPT, hygromycin phosphotransferase; NPTII, neomycin phosphotransferase II; kb, kilo-base pairs; Mb, mega-base pairs; GUS, b-glucuronidase; T-DNA, transferred DNA; I-*Sce*I, endonuclease I-*Sce*I

For use of the *Ac* element in gene-tagging, an understanding of the pattern of its transposition on a chromosome is clearly important. Previous genetic investigations in various laboratories have shown that *Ac* is preferentially transposed to a region within the same chromosome [5, 9, 10-12, 14, 22, 25, 30]. Insertional mutation of genes near the chromosomal location of *Ac* have, in fact, been obtained and the tagged genes have been cloned [8, 13, 15, 32]. Preferential transposition of *Drosophila P* elements to nearby chromosomal sites has also been reported [28]. However, our knowledge of distances and directions of transposition of *Ac* and other elements at the molecular level is very limited, and no systematic molecular analysis has been performed. In particular, it has yet to be solved whether *Ac* can transpose symmetrically or asymmetrically from a donor site on the same chromosome since results of genetic analyses with maize in two laboratories were inconsistent [9, 11]. Recently, we have reported molecular distances and directions on chromosomes of transposition of the *Ac* element in an Arabidopsis plant [18]. In the present lecture, we present the pattern of transposition of this element and discuss the feasibility of tagging of the nearest genes.

Experimental Design for Measurement of Distances of Transposition with I-*Sce*I

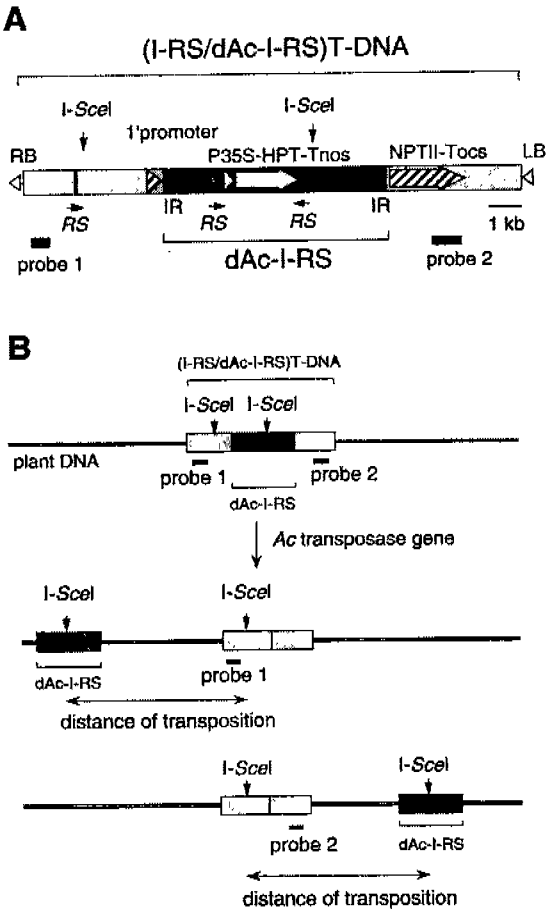

Figure 1. (A) Schematic diagram of the (I-RS/dAc-I-RS)T-DNA construct. dAc-I-RS indicates the non-autonomous derivative of the *Ac* element that contained I-*Sce*I sites and *RS*s (see text). (I-RS/dAc-I-RS)T-DNA indicates the T-DNA that containes a site for I-*Sce*I, *RS* and the dAc-I-RS element. LB and RB indicate left and right border sequences on the T-DNA, respectively. I-*Sce*I, the cleavage site for endonuclease I-*Sce*I; 1' promoter, the 1' promoter of the octopine Ti plasmid TR-DNA; P35S, the promoter of the gene for 35S RNA from cauliflower mosaic virus; HPT, the coding sequence of the gene for hygromycin phosphotransferase; NPTII, the coding sequence of the gene for neomycin phosphotransferase II; Tnos, the terminator of the gene for nopaline synthase; Tocs, the terminator of the gene for octopine synthase; IR, the terminal inverted repeat sequences of *Ac*; *RS*, the recombination site that is recognized by the R protein from *Zygosaccharomyces rouxii*. Thick lines indicate the probes (probes 1 and 2) used for Southern blotting analysis. (B) The strategy for measurement of the distance of transposition. I-*Sce*I indicates the site of cleavage by endonuclease I-*Sce*I. A black box indicates dAc-I-RS. Dotted boxes indicate regions of T-DNA other than dAc-I-RS in (I-RS/dAc-I-RS)T-DNA.

We investigated the molecular patterns of *Ac* transposition on the chromosomes of *Arabidopsis* using endonuclease I-*Sce*I. Since I-*Sce*I recognizes a specific 18-base-pair sequence [6, 7], the expected frequency of occurrence of the cleavage site for this enzyme is less than one for the chromosomes of *Arabidopsis* and, therefore, if such cleavage sites could be introduced into the *Ac* element and at the original chromosomal location of the *Ac* element, the physical distance of transposition could be directly determined by measuring the size of the DNA segment generated by digestion with I-*Sce*I of the genomic DNA.

Figure 1A shows the T-DNA construct that we made for the determination of distances of transposition. The T-DNA construct had two recognition sites for I-*Sce*I: one was inside and the other was outside the modified *Ac* transposable element We designated this element dAc-I-RS since it had the defective *Ac* element that contained the site for I-*Sce*I and *RS*s, which are recognition sites for a recombinase (R protein) in the R-*RS* recombination system from *Zygosaccaromyces rouxii* [1, 2, 16, 19, 20], though details of experimental results with this system will be described elsewhere. We also designated this T-DNA as (I-RS/dAc-I-RS) T-DNA since the T-DNA construct contained the site for I-*Sce*I, the *RS* sequence, and the dAc-I-RS element. After transposition of dAc-I-RS on the same chromosome, digestion of genomic DNA with I-*Sce*I should give rise a segment of chromosomal DNA flanked by part of the T-DNA at the original integration site and part of the transposed dAc-I-RS sequence (see Figure 1B). Using two DNA fragments as probes (probes 1 and 2) for Southern hybridization, as shown in Figure 1A and 1B, we were able to distinguish the relative direction of each transposition.

In addition to such specific DNA sequences, we inserted the gene for hygromycin phosphotransferase (HPT) under the control of P35S promoter into the dAc-I-RS element. For monitoring the excision of dAc-I-RS element , this element was inserted into the region between the 1' promoter [31] and the coding sequence of the gene for neomycin phosphotransferase II (NPTII) to generate a cryptic kanamycin resistance gene [3]. Thus, excision of dAc-I-RS in transgenic *Arabidopsis* plants that contain (I-RS/dAc-I-RS) T-DNA should create an active NPTII gene.

Measurement of Physical Distances of Transposition

We selected transgenic Arabidopsis plants that contained (I-RS/dAc-I-RS) T-DNA, and we identified plants, designated #14, #24 and #246, that had a single copy of the (I-RS/dAc-I-RS) T-DNA. The insertion site of (I-RS/dAc-I-RS) T-DNA in line #14 was mapped to position 80 on chromosome 1 by linkage analysis using tester lines (see Materials and Methods). The (I-RS/dAc-I-RS) T-DNAs in lines #24 and #246 were located at the top of chromosome 1 and the bottom of chromosome 2, respectively (Figure 2).

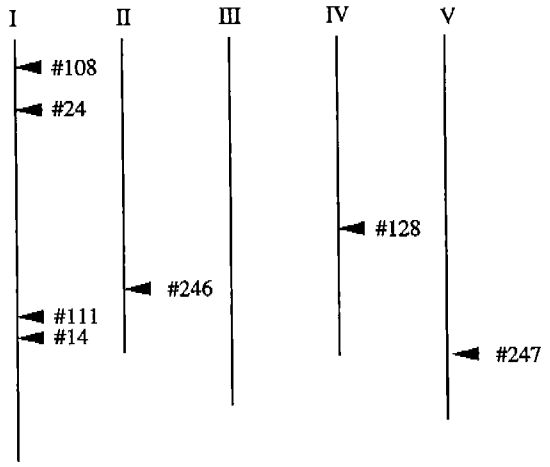

Figure 2. Chromosomal positions of integrated (I-RS/dAc-I-RS) T-DNAs in the transgenic Arabidopsis lines. Determination of integration sites in other 100 lines are in progress.

#246 were located at the top of chromosome 1 and the bottom of chromosome 2, respectively (Figure 2).

We crossed these transgenic plants with a transgenic plant (designated P35SAcTPase#9) that carried the gene for *Ac* transposase under the control of the CaMV 35S promoter (P35S). F1 plants were self-pollinated, and the resulting F2 plants were selected for kanamycin resistance to obtain plants in which the dAc-I-RS element presumably had excised from its original position. Such plants were also obtained from F3 progeny by the similar selection. We carried out Southern blotting analysis with genomic DNAs from the F2 and F3 plants to examine the reinsertion of dAc-I-RS. Two-thirds of the F2 and F3 progeny carried transposed dAc-I-RS elements.

High-molecular-weight chromosomal DNAs were purified from 60, 34 and 9 independent plants with a transposed dAc-I-RS that originated from transgenic lines #14, #24 and #246, respectively. The DNAs were digested with I-*Sce*I. The digests were fractionated by pulse-field gel electrophoresis (PFGE) and subjected to Southern blotting analysis with probes 1 and 2 depicted in Figure 1.

We analyzed a total of 60 transposition events from line #14 in this way. As summarized in Figure 3A, the dAc-I-RS element in 31 plants was transposed within a distance of 1,700 kb. In 21 out of 60 plants, the element was transposed to sites relatively close to the original site (within 200 kb).

In 29 out of the 60 plants, we detected no fragments of less than 1,700 kb by PFGE under our conditions. We carried out linkage analyses of the transposed dAc-I-RS element (hygromycin resistance) to the kanamycin-resistance marker at the original site in these plants (see Figure 1). The results of such analyses, together with the results of PFGE, demonstrate that overall, in about 75% of the events studied here, the dAc-I-RS element had transposed to sites that were genetically linked to the original site.

When we analyzed distances of transposition in progeny of transgenic lines #24 and #246, we obtained results similar to those in line #14 (data not shown).

We examined the orientation of the transposed dAc-I-RS elements at the insertion sites by Southern hybridization with probes that contained sequences specific for the left or the right side of the I-*Sce*I site in dAc-I-RS. These results of these experiments showed that the dAc-I-RS element could be inserted in both either orientation (data not shown).

The present result can provide a useful information for gene taqging by an *Ac-Ds* system. The short range transposition, preferentially within 200 kb, suggests that the *Ac* element might be useful for disrupting genes that are present in such adjacent regions. If *Arabidopsis* contains 20,000 genes, the average length of a single gene must be about 5 kb. The probability of direct tagging by the *Ac* element of a gene that is present between 0 and 200 kb can be estimated to be 99% on an average if we are to screen randomly isolated 1,000 plants that contain transposed elements.

Linkage of Transposed Elements to Known Loci

In order to relate "rightward" and "leftward" for T-DNA insert #14 to the genetic map of *Arabidopsis*, we made 3 point crosses of the NPTII and HPT in #14-58.3.1 and in #14-68.1.4 with flanking genetic loci *as2* (*asymmetry leaves 2*) and *cer5* . We found that the *as2* locus was present between the original position of the T-DNA and dAc-I-RS in #14-68.1.4, and that dAc-I-RS in #14-58.3.1 was located between the original position of the T-DNA and *cer5* (Figure 2B). These results also showed that dAc-I-RS elements in #14-58.3.1 and in #14-68.1.4 were transposed toward the centromere and the telomere at the bottom on the chromosome 1, respectively.

Isolation of Morphological Mutants with a Transposed dAc-I-RS

We searched for mutants with abnormal morphology among 103 plants that had a transposed dAc-I-RS from line #14 and found five mutants. Linkage analysis revealed that two out of the five mutations were closely linked to the transposed dAc-I-RS. One of these two mutations caused various abnormalities in leaf shape (named *abnormal leaf shape 1*: *ale1*) and

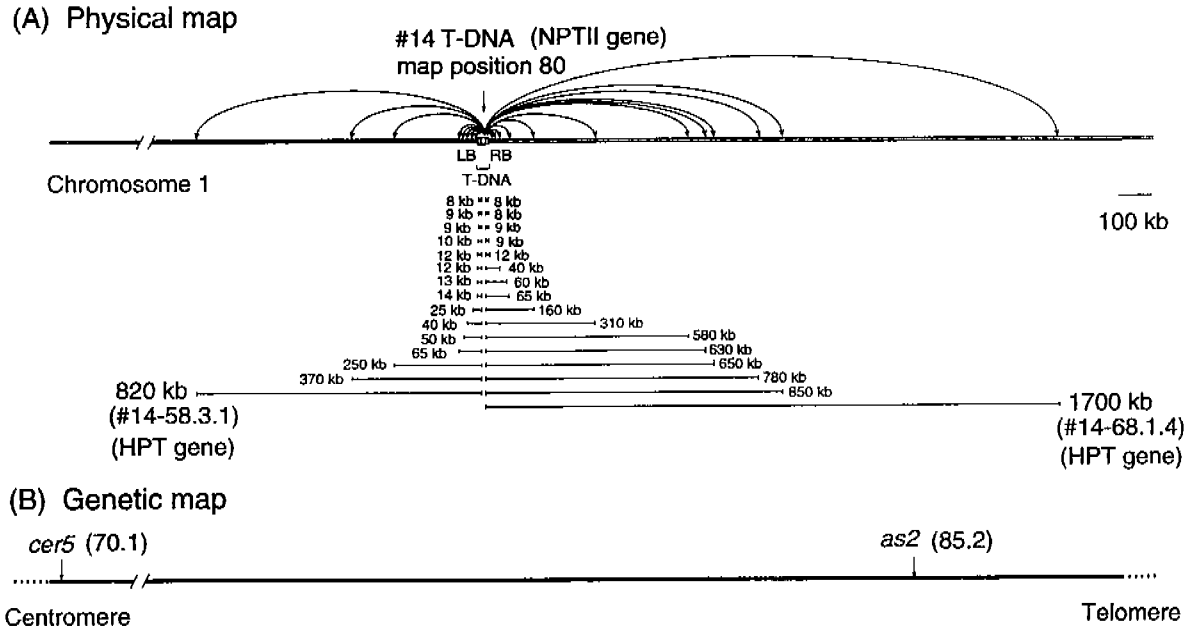

Figure 3. (A) Summary of distances of transposition in the I-RS/dAc-I-RS#14 line. LB and RB indicate left and right border sequences on the T-DNA, respectively. Each bar represents the distance of transposition in each plant that had a transposed dAc-I-RS. Some of the names of F3 progeny plants are indicated in parentheses. (B) Alignment of the genetic map around position 80. Relative positions of the transposed dAc-I-RS elements in #14-58.3.1 and #14-68.1.4 to *as2* and *cer5* were determined by three-point test and Southern blotting as described in the text.

the other caused delayed development and slow growth (*slow growth*: *slw1*). These two mutants had dAc-I-RS at a distance of 160 kb and 850 kb from the original site, respectively. The remaining three mutations did not cosegragate with the transposed dAc-I-RS.

We analyzed a phenotype of the *ale1* mutant, and isolated revertants. The mutant produced crinckled leaves with short petioles. Such abnormalities were observed mainly in juvenile leaves of the mutant but rarely in adult leaves. Microscopic analyses of the leaf sections and the leaf surfaces showed that severe abnormalities were found in the epidermis of cotyledons and juvenile leaves of *ale1*. Although pavement cells of the leaf surfaces of a wild-type Arabidopsis plant exhibited a unfixed- and rugged-shape like each piece of a jigsaw puzzle, those of the *ale1* mutant showed a much more smooth shape. In addition, organ adhesion was observed between leaves, between a cotyledon and a juvenile leaf, and betrween a stem and a juvenile leaf of the mutant. These observations indicate that the *ale1* epidermis was altered by a mutation.

We isolated revertants of the *ale1* mutant. In all the revertants, the dAc-I-RS elements were excised, and footprints of three or six nucleotides remained. These results reveal that the insertion of dAc-I-RS causes the *ale1* mutation. We cloned an *ALE1* gene and analzed characteristic features of the predicted ALE1 protein, which indicates that this gene is predicted to encode for a subtiline-like protease. We also analyzed sites of transcription of this gene. Based on the results of these experiments, we will propose a model how differentiation of plant epidermal cells can be controlled during the development.

In the *slw* mutant, the dAc-I-RS was inserted in the NADPH reductase gene [27]. Although the mutation was not linked to the insertion, its site was mapped in the cjromosomal site very close to the NADPH reductase gene. Identification of the mutated gene in *slw* is in progress in my laborartory.

Use of the Transposed dAc-I-RS for Map-Based Cloning of Genes

Linkage analysis of previously mapped genes, such as *as2* and transposed dAc-I-RS elements with hygromycin resistance, together with measurements of physical distances of transposition allowed us to determine the precise positions of these genes and of the dAc-I-RS elements on the chromosomal DNA. In addition, since the transposition site and the original site of the T-DNA include the HPT and NPTII genes, respectively, it should now be easy to obtain recombinants between these resistance genes and the mutant genes. Recently, we have cloned the *AS2* gene

by using such a strategy. In the present lecture, we also provide several pieces of information of symmetrical development of plant leaves and a structure of the predicted AS2 protein.

Use of the Transposed dAc-I-RS in Combination with the R-RS Site-Specific Recombination System

Development of recombinant DNA technology for manipulation of large segments of DNA is of importance for the structural and functional analysis of eukaryotic chromosomes. For development of such a technology in a plant, use of site-specific recombination systems, such as the cre-*loxP* system, the R-*RS* system and FLP-*FRT* systems, has been proposed [17, 21, 23, 24, 26, 29]. Chromosomal deletions and inversions have been induced in *Arabidopsis* and tobacco plants by exploitation of these systems [21, 26]. However, full exploitation of this system for systematic analysis has been cumbersome, due in part to technical difficulties. One of the aims of our present study was to solve this problem. As shown in Figure 1, the (I-RS/dAc-I-RS) T-DNA that we constructed contained two *RS* sequences, namely, recognition sites for a recombinase (R protein) from *Z. rouxii*. The *Arabidopsis* plants that we generated in the present study will, therefore, be useful for a systematic examination of chromosomal deletions mediated by the R-*RS* system since the physical distance between two *RS* sequences on the chromosome have been precisely determined.

Acknowledgments

CM was supported by a grant for Precursory Research for Embryonic Science and Technology from the Research Development Corporation of Japan. HT and SH were supported by JSPS Research Fellowships for Young Scientists. This work was also supported in part by Grants-in-Aid for Scientific Research on Priority Areas and for General Scientific Research on Priority Areas from the Ministry of Education, Science and Culture of Japan (No. 06278103), and by a grant for Pioneering Research Projects in Biotechnology from the Ministry of Agriculture, Forestry, and Fisheries of Japan.

References

1. Araki H, Jearnpipatkul A, Tatsumi H, Sakurai T, Ushio K, Muta T, Oshima Y: J Mol Biol 182: 191-203 (1985).
2. Araki H, Nakanishi N, Evans BR, Matsuzaki H, Jayaram M, Oshima Y: J Mol Biol 225: 25-37 (1992).
3. Baker B, Coupland G, Fedoroff N, Starlinger P, Schell J: EMBO J 6: 1547-1554 (1987).
4. Bancroft I, Dean C: Genetics 134: 1221-1229 (1993).
5. Carrol BJ, Klimyuk VI., Thomas CM, Bishop GJ, Harrison K, Scofield SR, Jones JD: G Genetics 139: 407-420 (1995).
6. Colleaux L, d'Auriol L, Betermier M, Cottarel G, Jacquier A, Galibert F, Dujon B: Cell 44: 521-33 (1986).
7. Colleaux L, d'Auriol L, Galibert F, Dujon B: Proc Natl Acad Sci USA 85: 6022-6026 (1988).
8. Delong A, Calderon-Urrea A, Dellaporta SL: Cell 74: 757-768 (1993).
9. Dooner HK, Belachew A: Genetics 122: 447-457 (1989).
10. Dooner HK, Keller J, Haper E, Ralston E: Plant Cell 3: 473-482 (1991).
11. Greenblatt IM: Genetics 108: 471-485 (1984).
12. Greenblatt IM: Brink RA: Genetics 47: 489-501 (1962).
13. James Jr DW, Lim E, Keller J, Plooy I, Ralston E, Dooner HK: Plant Cell 7: 309-319 (1995).
14. Jones JDG, Carland F, Lim E, Ralston E, Dooner HK: Plant Cell 2: 701-707 (1990).
15. Jones DA, Thomas CM, Hammond-Kosack KE, Balint-Kurti PJ, Jones JDG: Science 266: 789-793 (1994).
16. Kawasaki H, Matsuzaki H, Nakajima R, Oshima Y: Yeast 7: 859-865 (1991).
17. Machida C, Onouchi H, Machida Y: Nippon Nougeikagaku Kaishi 67: 711-715 (1993).
18. Machida C, Onouchi H, Koizumi J, Hamada S, Semiarti E, Torikai S, Machida Y: Proc Natl Acad Sci USA 94: 8675-8680 (1997).
19. Matsuzaki H, Araki H, Oshima Y: Mol Cell Biol 8: 955-962 (1988).
20. Matsuzaki H, Nakajima R, Nishiyama J, Araki H, Oshima Y: J Bacteriol 172: 610-618 (1990).
21. Medberry SL, Dale E, Qin M, Ow DW: Nucleic Acids Res 23: 485-490 (1995).
22. Moreno M, Chen J, Greenblatt I, Dellaporta SL: Genetics 131: 939-956 (1992).
23. Onouchi H, Yokoi K, Machida C, Matsuzaki H, Ohshima Y, Matsuoka M, Nakamura K, Machida Y: Nucleic Acids Res 19: 6373-6378 (1991).
24. Onouchi H, Nishihama R, Kudo M, Machida Y, Machida C: Mol Gen Genet 247: 653-660 (1995).
25. Osborne BI, Corr CA, Prince JP, Hehl R, Tanksley SD, McCormick S, Baker B: Genetics 129: 833-844 (1991).
26. Osborne BI, Wirtz U, Baker B: Plant J 7: 687-701 (1995).
27. Torres MA, Onouchi H, Hamada S, Machida C, Hammond-Kosack KE, Jones JDG: The Plant J 14: 365-370 (1998).
28. Tower J, Karpen GK, Craig N, Spradling AC: Genetics 133: 347-359 (1993).
29. van Haaren JJ, Ow DW Plant Mol Biol 23: 525-533 (1993).
30. van Schaik N, Brink RA: Genetics 44: 725-738 (1959).
31. Velten J, Velten L, Hain R, Schell J: EMBO J 3: 2723-2730 (1984).
32. Whitham S, Dinesh-Kumar SP, Choi D, Hehl R, Corr C, Baker B: Cell 78: 1-20 (1994).