

# Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*

Asuka Miura\*†, Shoji Yonebayashi†‡, Koichi Watanabe§, Tomoko Toyama§, Hiroaki Shimada|| & Tetsuji Kakutani\*‡§

\* National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan

‡ National Institute of Agrobiological Resources, Tsukuba, Ibaraki 305-8602, Japan

§ CREST, Japan Science and Technology Corporation, Tokyo 101-0062, Japan

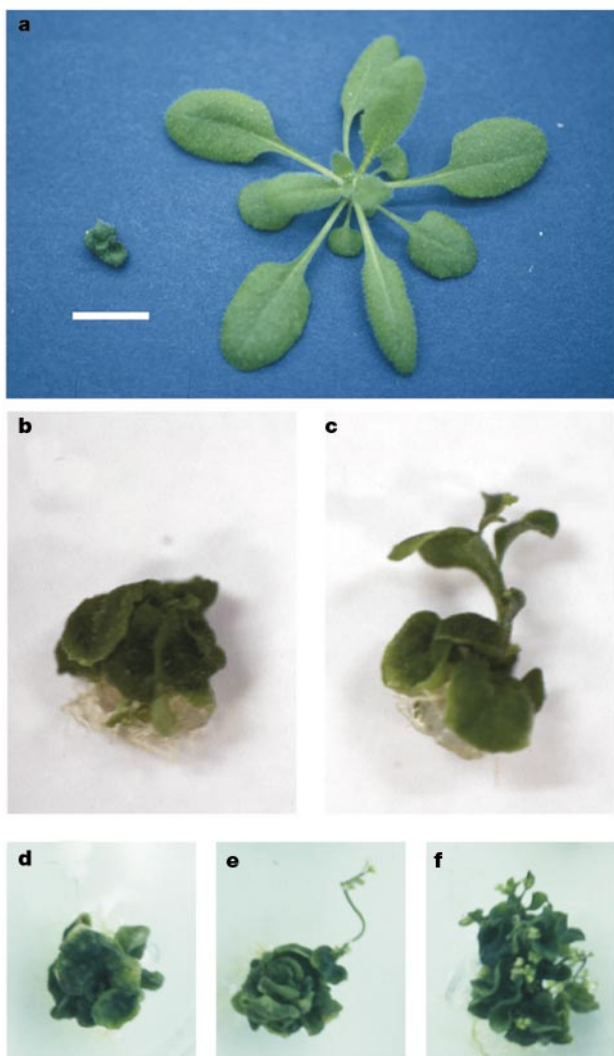
|| Science University of Tokyo, Noda, Chiba 278-8510, Japan

† These authors contributed equally to this work.

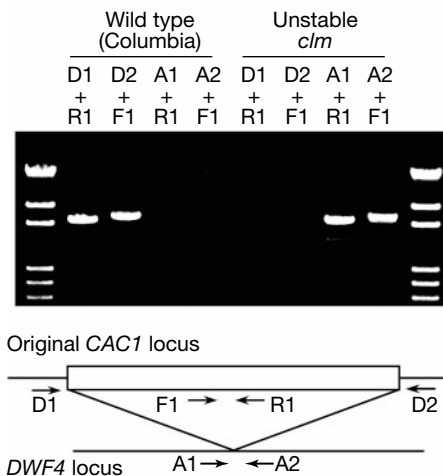
A major component of the large genomes of higher plants and vertebrates comprises transposable elements and their derivatives, which potentially reduce the stability of the genome<sup>1</sup>. It has been proposed that methylation of cytosine residues may

suppress transposition, but experimental evidence for this has been limited<sup>2-5</sup>. Reduced methylation of repeat sequences results from mutations in the *Arabidopsis* gene *DDM1* (*decrease in DNA methylation*)<sup>6</sup>, which encodes a protein similar to the chromatin-remodelling factor SWI2/SNF2 (ref. 7). In the *ddm1*-induced hypomethylation background, silent repeat sequences are often reactivated transcriptionally, but no transposition of endogenous elements has been observed<sup>8-11</sup>. A striking feature of the *ddm1* mutation is that it induces developmental abnormalities by causing heritable changes in other loci<sup>12,13</sup>. Here we report that one of the *ddm1*-induced abnormalities is caused by insertion of *CAC1*, an endogenous CACTA family transposon. This class of *Arabidopsis* elements transposes and increases in copy number at high frequencies specifically in the *ddm1* hypomethylation background. Thus the *DDM1* gene not only epigenetically ensures proper gene expression<sup>13-16</sup>, but also stabilizes transposon behaviour, possibly through chromatin remodelling or DNA methylation.

To study the molecular basis for *ddm1*-induced developmental abnormalities, we identified the mutated gene responsible for one of them, *clam* (*clm*), which is characterized by lack of elongation in shoots and petioles<sup>12</sup> (Fig. 1a). This phenotype was initially unstable, and phenotypically normal sectors were occasionally observed (Fig. 1b-f). In subsequent generations, the *clm* phenotype stabilized in some of the progeny families and was inherited as a recessive mendelian trait, which could be mapped genetically<sup>12</sup>. By genotyping 926 chromosomes from a mapping cross, we narrowed the *clm* locus to a 64-kilobase (kb) region in bacterial artificial chromosome (BAC) clone T3A5 (GenBank AL132979) on chromosome 3. This region contains the *DWF4* gene, encoding 22- $\alpha$ -hydroxylase, a 513-amino-acid protein that mediates the biosynthesis of brassinosteroid, a regulator of cell elongation<sup>17</sup>. Complementation tests indicated that *clm* is allelic to the *dwf4-4* mutation (see Supplementary Information). The sequence of the *DWF4* gene in the stable *clm* plant revealed a 4-base-pair (bp) insertion mutation in the second exon at nucleotide +527 from the translation start site. This insertion converts TAG to TAGC-TAG, creating a stop codon after the 149th amino acid. Although this defect is probably the cause of the stable *clm* phenotype, it cannot account for the instability of the *clm* phenotype in initial generations.

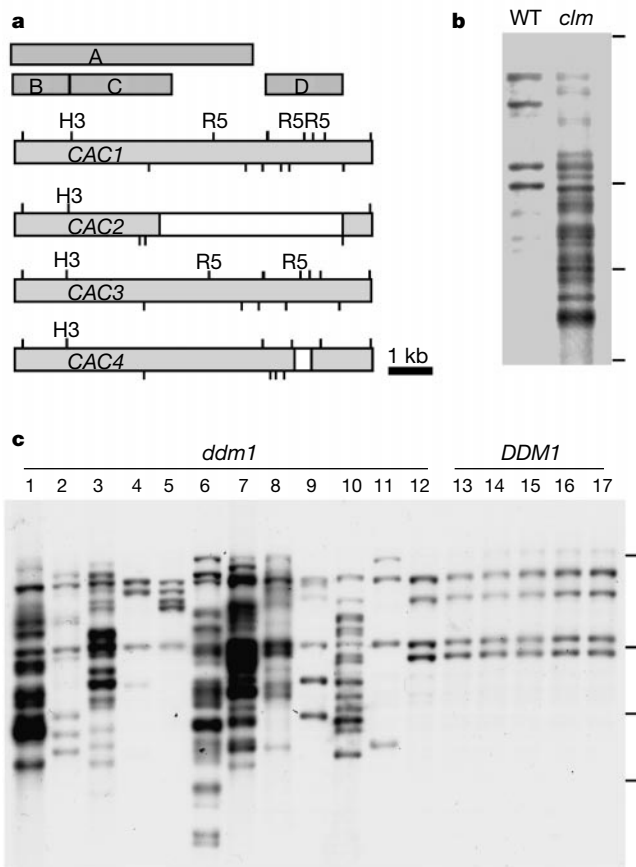


**Figure 1** Stable and unstable *clm* phenotypes. **a**, A wild-type Columbia plant showing expanded leaves (right) and a *clm* homozygous plant with a defect in leaf and stem elongation (left). Both are 35 days old. **b-f**, *clm* homozygous plants with (**c, e, f**) or without (**b, d**) reversion sectors. Plants were grown on soil (**a**) or nutrient agar plates<sup>12</sup> (**b-f**). Scale bar, 10 mm.

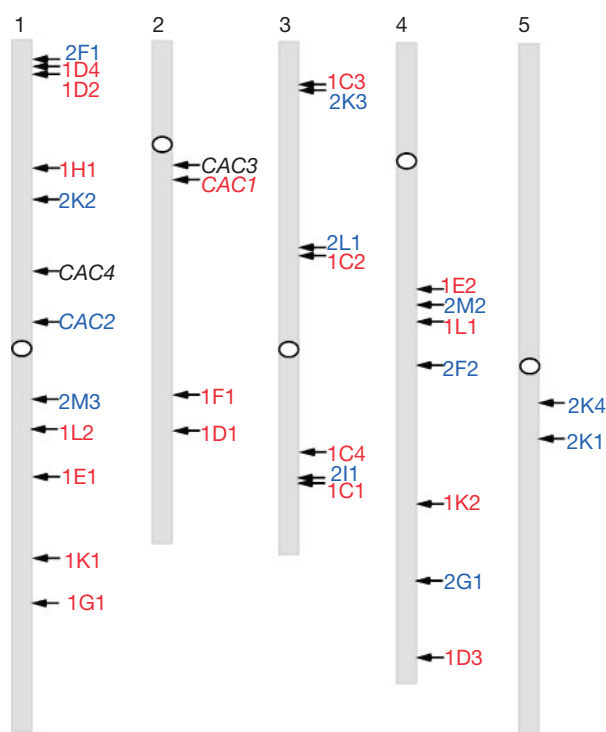


**Figure 2** Transposition of *CAC1* element from the original donor site to the *DWF4* locus. Arrows indicate primers used for PCR. The DNA length markers are 19.3, 7.74, 5.53, 3.14, 2.69 and 2.32 kb. The sizes of transposon and the starting position (but not the length) of arrows for primers reflect the physical length.

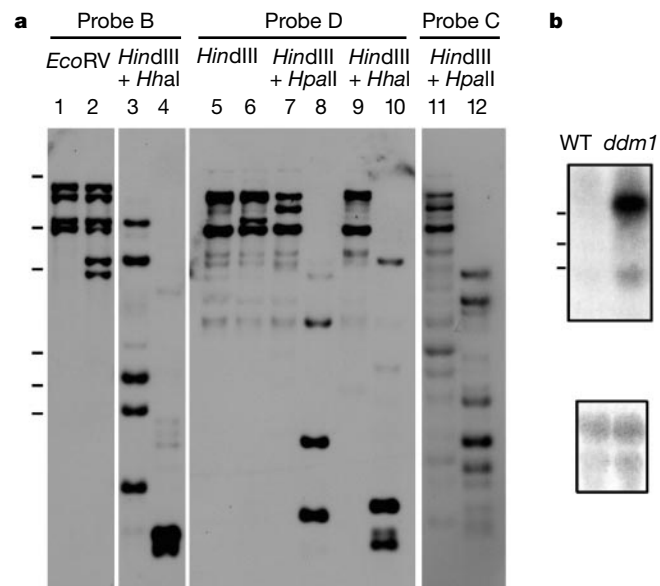
In the unstable *clm* plant, we used genomic Southern analysis to identify a several-kilobase insertion in the *DWF4* gene (data not shown). We amplified the inserted DNA from unstable *clm* plants using the suppression polymerase chain reaction (PCR) technique<sup>18</sup>. Sequences obtained from each side of the insertion (523 and 154 nucleotides) showed exact matches to the 3' and 5' ends of a 8,479-bp sequence in chromosome 2 (nucleotides 52,296 to 60,774 in GenBank AC005897). No other exact match was found in the *Arabidopsis* genome database (<http://www.arabidopsis.org/blast/>). This 8,479 sequence has features typical of the CACTA family of transposons, which include maize *Spm/En* and snapdragon *Tam1* elements: the 5' and 3' ends have the conserved terminal inverted sequence CACTACAA, and the internal region includes a predicted gene (GenBank AAC97237) with similarity to the putative snapdragon transposase Tnp2 (ref. 19). To confirm the transposition of this sequence, we amplified two DNA fragments covering the entire element using internal sequences and the *DWF4* gene sequence as primers (Fig. 2). The estimated length of the product and the restriction digestion pattern were identical for the original copy and the insert in the *DWF4* gene, suggesting that the full-length element had transposed from chromosome 2 to the *DWF4* gene on chromosome 3. This transposon, designated *CACTA1* (*CAC1*), was responsible for the unstable *clm* phenotype;



**Figure 3** Transposition of *CAC* elements in the *ddm1* mutant background. **a**, The *CAC* transposon family. Open boxes in *CAC2* and *CAC4* show the internal deletions. Vertical bars show the restriction sites: H3, *Hind*III; R5, *Eco*RV; other bars above the elements, *Hha*I; below the elements, *Hpa*I. The transpositions of *CAC* elements were examined by Southern blot analysis using *Eco*RV and probe A. **b**, *clm* and wild-type Columbia (WT) plants. **c**, The *ddm1* and wild-type *DDM1* lines self-pollinated six to seven times in parallel<sup>12</sup>. Five out of 12 *DDM1* lines are shown. The DNA length markers are 19.3, 7.74, 5.53 and 4.25 kb.



**Figure 4** Transposition of *CAC1* and *CAC2* elements to unlinked loci. Sequences flanking the transposed copies were determined by suppression PCR (for determined sequences and methods see Supplementary Information). The red and blue loci show *CAC1* and *CAC2* copies, respectively. The transposed copies with second character D, E, F, G, H, I, K, L, M and C are from the *ddm1* lines shown in lanes 10, 9, 8, 7, 6, 5, 3, 2 and 1 of Fig. 3c and the *clm* line (Fig. 3b), respectively.



**Figure 5** The *ddm1*-specific hypomethylation and RNA accumulation of *CAC* elements. **a**, Methylation-sensitive restriction enzymes (*Hpa*I or *Hha*I) and probes B, C, D (Fig. 3a) were used to compare the methylation status of *CAC* elements between *ddm1* (even lanes) and Columbia wild-type (odd lanes) plants. The *ddm1* plant is before the repeated self-pollination (four generations before the plant shown in lane 10 of Fig. 3c). It still keeps the donor copies of *CAC* elements (lane 2). The DNA length markers are 19.3, 7.74, 5.53, 3.14, 2.69 and 2.32 kb. **b**, RNA blot analysis. Probe A (Fig. 3a) was used to detect *CAC* transcript in wild-type and *ddm1* (*clm*) plants. The RNA length markers are 6, 4 and 3 kb. Bottom panel, ribosomal RNA on the filter stained with methylene blue.

in the sector without the *clm* phenotype, restoration of the insertion site to the normal structure was found by sequencing PCR-amplified templates (see Supplementary Information).

The *Arabidopsis* genome contains three additional sequences distinct from but significantly similar to *CAC1*; their terminal 200 bp are 90–98% identical (Fig. 3a). As they are likely to comprise a new transposon family with *CAC1*, we designated them *CAC2* (GenBank AC069160, nucleotides 34,626–38,790), *CAC3* (GenBank AC006429, nucleotides 85,404–76,949) and *CAC4* (GenBank AC027135, nucleotides 72,903–64,858). Genomic Southern analysis using the methylation-insensitive restriction enzyme *EcoRV* revealed that *CAC* elements transposed and increased in the copy number in the *clm* line (Fig. 3b).

To see whether the mobilization of the *CAC* elements is a general phenomenon in the *ddm1* background, we examined genomic DNA of twelve *ddm1* lines that had independently and randomly self-pollinated six to seven times<sup>12</sup> (Fig. 3c). Eleven out of twelve *ddm1* lines exhibited changes in the band pattern of *CAC* elements, and six showed several-fold increases in band number (up to >20). Transposition of *CAC1* and *CAC2* was confirmed by sequencing the genomic regions flanking them. Both were found to transpose to unlinked loci throughout the genome (Fig. 4). In contrast, we never observed the change in *EcoRV* banding pattern for the *CAC* elements in any of the 12 control wild-type *DDM1* lines that had independently self-pollinated seven times in parallel (Fig. 3c), indicating that the mobilization of *CAC* elements is a consequence of the *ddm1* mutation.

Mobilization of transposons is often associated with hypomethylation and transcriptional activation<sup>20–23</sup>. Indeed, the *CAC* elements were hypomethylated and transcriptionally activated in *ddm1* plants (Fig. 5), like other repeated sequences<sup>8–11</sup>. PCR with reverse transcriptase and sequencing confirmed that at least the *CAC1* transcript was highly accumulated in *ddm1* (unpublished data). However, it remains to be seen whether transcriptional activation alone is sufficient to trigger the *ddm1*-induced transpositional activation of *CAC* elements; the *ddm1* mutation might also affect the epigenetic chromatin state, as the gene encodes a protein similar to the chromatin-remodelling factor SWI2/SNF2.

Maize transposons were the first plant genes found to be epigenetically regulated<sup>24,25</sup>, and the epigenetic control of the maize *Spm* by *cis* elements and transposon-encoded Tnp proteins have been studied extensively<sup>21</sup>. Little is known, however, about controlling host factors. We report here endogenous *Arabidopsis* elements that are mobilized by a host gene mutation affecting epigenetic states. In addition to *ddm1*, a variety of *Arabidopsis* mutants in gene silencing and methylation<sup>26–32</sup> provide excellent systems to dissect mechanistically the transpositional activation of *CAC* elements, which should further clarify the role of epigenetic controls on genome evolution.

**Note added in proof:** It has been shown recently that *Arabidopsis* elements similar to Robertson's *Mutator* in maize become active in the *ddm1* mutant<sup>33</sup>. □

## Methods

The growth conditions, preparation of plant genomic DNA and mapping of the *CLM* locus have been described<sup>12</sup>. The suppression PCR method was as described<sup>18</sup>. We used *EcoRV*-digested genomic DNA for adapter ligation and subsequent amplification. The sequences of primers used for mapping the *CLM* locus, sequencing the *DWF4* gene, suppression PCR reactions, transposition monitoring (Fig. 2) and probe amplification (Fig. 3) are listed as Supplementary Information or on <http://www.nig.ac.jp/labs/AgriGen/cacta-Nature.html>. The Supplementary Information also includes sequences flanking the transposed *CAC* copies (Fig. 4).

Received 11 January; accepted 21 February 2001.

1. SanMiguel, P. *et al.* Nested retrotransposon in the intergenic region of maize genome. *Science* **274**, 765–768 (1996).

- Yoder, J. A., Walsh, C. P. & Bestor, T. H. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13**, 335–340 (1997).
- Martienssen, R. Transposons, DNA methylation and gene control. *Trends Genet.* **14**, 263–264 (1998).
- Walsh, C. P., Chaillet, J. R. & Bestor, T. H. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nature Genet.* **20**, 116–117 (1998).
- Matzke, M. A., Mette, M. F., Aufsatz, W., Jakowitsch, J. & Matzke, A. J. Host defenses to parasitic sequences and the evolution of epigenetic control mechanisms. *Genetica* **107**, 271–287 (1999).
- Vongs, A., Kakutani, T., Martienssen, R. A. & Richards, E. J. *Arabidopsis thaliana* DNA methylation mutants. *Science* **260**, 1926–1928 (1993).
- Jeddeloh, J. A., Stokes, T. L. & Richards, E. J. Maintenance of genomic methylation requires a SWI2/SNF2-like protein. *Nature Genet.* **22**, 94–97 (1999).
- Mittelsten-Scheid, O., Afsar, K. & Paszkowski, J. Release of epigenetic gene silencing by trans-acting mutations in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **95**, 632–637 (1998).
- Jeddeloh, J. A., Bender, J. & Richards, E. J. The DNA methylation locus *DDM1* is required for maintenance of gene silencing in *Arabidopsis*. *Genes Dev.* **12**, 1714–1725 (1998).
- Steimer, A. *et al.* Endogenous targets of transcriptional gene silencing in *Arabidopsis*. *Plant Cell* **12**, 1165–1178 (2000).
- Hirochika, H., Okamoto, H. & Kakutani, T. Silencing of retrotransposons in *Arabidopsis* and reactivation by the *ddm1* mutation. *Plant Cell* **12**, 357–369 (2000).
- Kakutani, T., Jeddeloh, J., Flowers, S., Munakata, K. & Richards, E. Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. *Proc. Natl Acad. Sci. USA* **93**, 12406–12411 (1996).
- Kakutani, T. Genetic characterization of late-flowering traits induced by DNA hypomethylation mutation in *Arabidopsis thaliana*. *Plant J.* **12**, 1447–1451 (1997).
- Soppe, W. *et al.* The gain of function epi-mutant *FWA* causes late flowering. *Mol. Cell* **6**, 791–802 (2000).
- Jacobsen, S., Sakai, H., Finnegan, E., Cao, X. & Meyerowitz, E. Ectopic hypermethylation of flower-specific genes in *Arabidopsis*. *Curr. Biol.* **10**, 179–186 (2000).
- Kakutani, T., Munakata, K., Richards, E. & Hirochika, H. Meiotically and mitotically stable inheritance of DNA hypomethylation induced by *ddm1* mutation of *Arabidopsis thaliana*. *Genetics* **151**, 831–838 (1999).
- Choe, S. *et al.* The *DWF4* gene of *Arabidopsis* encodes a cytochrome P450 that mediates multiple 22 $\alpha$ -hydroxylation steps in brassinosteroid biosynthesis. *Plant Cell* **10**, 231–243 (1998).
- Siebert, P. D., Chenchik, A., Kellogg, D. E., Lukyanov, K. A. & Lukyanov, S. A. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res.* **23**, 1087–1088 (1995).
- Nacken, W. K., Piotrowski, R., Saedler, H. & Sommer, H. The transposable element *Tam1* from *Antirrhinum majus* shows structural homology to the maize transposon *En/Spm* and has no sequence specificity of insertion. *Mol. Gen. Genet.* **228**, 201–208 (1991).
- Chandler, V. L. & Walbot, V. DNA modification of a maize transposable element correlates with loss of activity. *Proc. Natl Acad. Sci. USA* **83**, 1767–1771 (1986).
- Fedoroff, N. in *Epigenetic Mechanisms of Gene Regulation* (eds Riggs, A., Martienssen, R. & Russo, V.) 575–592 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1996).
- Scortecci, K. C., Dessaux, Y., Petit, A. & Van Sluys, M. A. Somatic excision of the *Ac* transposable element in transgenic *Arabidopsis thaliana* after 5-azacytidine treatment. *Plant Cell Physiol.* **38**, 336–343 (1997).
- Lawson, E. J., Scofield, S. R., Sjodin, C., Jones, J. D. & Dean, C. Modification of the 5' untranslated leader region of the maize Activator element leads to increased activity in *Arabidopsis*. *Mol. Gen. Genet.* **245**, 608–615 (1994).
- McClintock, B. Chromosome organization and genic expression. *Cold Spring Harbor Symp. Quant. Biol.* **16**, 13–47 (1951).
- McClintock, B. The suppressor-mutator system of control of gene action in maize. *Carnegie Institution of Washington Year Book* **57**, 415–429 (1958).
- Finnegan, E., Peacock, J. & Dennis, E. Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. *Proc. Natl Acad. Sci. USA* **93**, 8449–8454 (1996).
- Ronemus, M., Galbiati, C., Ticknor, J., Chen, J. & Dellaporta, S. Demethylation-induced developmental pleiotropy in *Arabidopsis*. *Science* **273**, 654–657 (1996).
- Furner, I. J., Sheikh, M. A. & Collett, C. E. Gene silencing and homology-dependent gene silencing in *Arabidopsis*. Genetic modifiers and DNA methylation. *Genetics* **149**, 651–662 (1998).
- Amedeo, P., Habu, Y., Afsar, K., Scheid, O. M. & Paszkowski, J. Disruption of the plant gene *MOM* releases transcriptional silencing of methylated genes. *Nature* **405**, 203–206 (2000).
- Mourrain, P. *et al.* *Arabidopsis* *SGS2* and *SGS3* genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell* **101**, 533–542 (2000).
- Dalmay, T., Hamilton, A., Rudd, S., Angell, S. & Baulcombe, D. C. An RNA-dependent RNA polymerase gene in *Arabidopsis* is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell* **101**, 543–553 (2000).
- Fagard, M. *et al.* *AGO1*, *QDE-2*, and *RDE-1* are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. *Proc. Natl Acad. Sci. USA* **97**, 11650–11654 (2000).
- Singer, T., Yordan, C. & Martienssen, R. A. Robertson's *Mutator* transposons in *A. thaliana* are regulated by the chromatin-remodeling gene *Decrease in DNA Methylation (DDM1)*. *Genes Dev.* **15**, 591–602 (2001).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

## Acknowledgements

We thank K. Munakata for technical assistance, A. Miyao for advice on the suppression PCR technique, and R. Martienssen and Y. Hiromi for comments on the manuscript. We acknowledge ABRC at the Ohio State University for the BAC clones.

Correspondence and requests for material should be addressed to T.K. (e-mail: tkakutan@lab.nig.ac.jp). The *CAC1*, *CAC2*, *CAC3* and *CAC4* sequences are deposited in GenBank under accession numbers AB052792, AB052793, AB052794 and AB052795, respectively.