Cell. Mol. Life Sci. 62 (2005) 3080–3091
1420-682X/05/243080-12
DOI 10.1007/s00018-005-5301-2

Cellular and Molecular Life Sciences

Research Article

Transgene silencing may be mediated by aberrant sense promoter sequence transcripts generated from cryptic promoters

M. C. Eike[+,†], I. S. Mercy[++,†] and R. B. Aalen*

Department of Molecular Biosciences, University of Oslo, PO Box 1041 Blindern, 0316 Oslo (Norway),
Fax: + 47 22857297, e-mail: reidunn.aalen@imbv.uio.no

Received 6 July 2005; received after revision 4 October 2005; accepted 7 October 2005
Online First 28 November 2005

Abstract. To investigate whether silencing of a T-DNA-carried *nptII* gene in five single-copy transgenic Arabidopsis lines might be due to position effects, genomic DNA flanking the insertions was analysed for gene density, GC content, presence of short repeats and transposable elements, i.e. factors suggested to promote silencing. No single, common factor could explain the observed silencing. However, in two lines, a transcript covering the *nos* promoter driving the *nptII* gene was detected. In sibling sublines with ~100% silencing, the *nos* promoter was heavily methylated. *In silico* analysis suggested the presence of cryptic core promoters upstream of the *nos* promoter, in one case in the plant DNA and in the other in a short inverted T-DNA region. These fragments were able to drive reporter gene expression in transgenic *Arabidopsis* plants. Our results indicate that methylation and silencing of transgenic promoters may be mediated by aberrant RNA transcribed from cryptic promoters at the transgene insertion site.

Key words: Transgene silencing; aberrant RNA; DNA methylation; cryptic promoter.

Genetic modification of plants is important both for applied agricultural purposes and in research, where transgene studies represent powerful tools in identifying and studying gene functions. However, in some cases, transgenes are not expressed as expected. Transgene silencing can be triggered between two related transgenes, by RNA or DNA viruses with homologous sequences and between a transgene and a homologous endogenous gene, i.e. phenomena included under the broader term homology-dependent gene silencing (HDGS) [1, 2].

These phenomena work on at least two levels. In transcriptional gene silencing (TGS) [1], mRNA synthesis is blocked, while in post-transcriptional gene silencing (PTGS, analogous to RNA interference/RNAi) the RNA is degraded before it can be translated [reviewed in ref. 3]. PTGS may be triggered by transcript levels above a certain threshold [4]. Both TGS and PTGS/RNAi are often accompanied by changes in DNA methylation levels and/or chromatin structure. In recent years, HDGS has been explained by models proposing RNA-mediated mechanisms involving double-stranded RNA (dsRNA) [reviewed in ref. 2].

The most direct way of producing dsRNA is by read-through transcription of inverted repeats, which are

* Corresponding author.
Present addresses:
[+] Institute of Immunology, Rikshospitalet University Hospital and University of Oslo, 0027 Oslo (Norway)
[++] The Norwegian School of Veterinarian Science, P.O. Box 8146 Dep 0033 Oslo (Norway)
[†] These authors contributed equally to this work.

strong inducers of PTGS or TGS of homologous sequences [5–7]. Alternatively, transcripts may serve as targets for complementary strand synthesis. Several models postulate the production of aberrant RNAs that serve as templates for RNA-dependent RNA polymerase (RdRP), resulting in silencing-inducing dsRNA [8, 9].

Reporter-gene expression levels from single-copy transgenes are often high and stable, and silencing is rare [4, 10]. Moreover, in the few reported cases of silencing of single-copy transgenes, the involvement of homology-dependent mechanisms is not obvious [10–15]. Conversely, the presence of repeated transgenic sequences will not always result in silencing [see e. g. ref. 14]. Therefore, other factors, either features of the inserted DNA or of the genomic DNA flanking foreign DNA, have been suggested to trigger silencing.

Position effect variegation (PEV) is a well-known phenomenon in *Drosophila*, and is characterised by stochastic inactivation of genes located in the vicinity of a heterochromatic region [reviewed in ref. 16]. Similarly, there are examples of stochastically inactivated transgenes inserted within or in the vicinity of centromeric and intercalary heterochromatin [17, 18] or telomeres [19]. These regions are frequently associated with a repeated DNA structure [reviewed in ref. 20]. Accordingly, the presence of repeats in flanking sequences has been associated with variegated transgene silencing in petunia, tobacco and barley [21–23] and paramutation in maize [24]. Remnants of retroelements have been identified in the vicinity of unstably expressed transgenes [18], and methylation of retroelements has been suggested to affect adjacent sequences, leading to transcriptional repression [25, 26]. Conversely, matrix attachment regions (MARs), dispersed throughout eukaryotic genomes [27] have in some cases been shown to confer position- and copy-number-independent control of transgene expression [28–30]. Furthermore, abrupt changes in GC content between genomic DNA and a transgenic insert have been suggested to make the transgene conspicuous to silencing mechanisms [31, 32]. In plants, the GC-rich binary vector backbone (BVB) sequences sometimes inserted in the genome together with transfer DNA (T-DNA) in *Agrobacterium*-mediated transformation, have been suggested to serve as a prominent target for DNA methyltransferases [17, 18].

Promoter/enhancer trap vectors have shown that promoterless transgenes can be expressed from endogenous as well as cryptic promoters, i. e. elements that are non-functional at their native positions in the genome, but functional when positioned adjacent to genes [33, 34]. This suggests that variable expression and even silencing of transgenes could be caused by interference from transcription initiated from adjacent promoters [35].

In a screen for *Arabidopsis thaliana* lines displaying silencing of a T-DNA-carried *nptII* gene normally conferring Kanamycin (Km) resistance, we have previously identified single-locus lines, including some with a single T-DNA insertion [14, 36]. In the present article we used *in silico* and molecular tools to elucidate factors that may be involved in silencing in a subset of these lines. As an initial approach, plant DNA flanking the borders of the T-DNAs was analysed for gene and GC content, presence of various repeats and MARs, as well as the theoretical possibility of transcriptional interference. However, *in silico* analysis alone may not be the best means of identifying position effects. Therefore, we conducted a detailed molecular analysis of two silenced lines in search of aberrant and antisense transcripts. In addition, genomic bisulphite sequencing was used to investigate DNA methylation in the *nos* promoter in sublines displaying silencing.

Materials and methods

Plant material and constructs. Plant material, growth conditions and Km selection have been described in Meza et al. [36]. Five silencing lines transformed with the constructs pKOH110 35SGUS (K lines) or pPCV002 35SGUS (P lines) (fig. 1) were used [14, 37], as well as three lines without silencing (line 200 *ida*, SENAPE4/775 and LIGUINE1/747) from a mutagenesis/promoter-trap project [38–40] that harbour the construct pMHA2 identical to pKOH110 35SGUS except for the lack of the 35S promoter (fig. 1A).

Cloning of flanking regions. Cloning of flanking regions has been described in Meza et al. [14]. However, in this work, the plant DNA flanking the right border of line P10 was cloned using the genomic primer 146LP 390 (5'-AAGCGTGACTACAATTCGGAAGC-3') and the T-DNA primers 146RP 1787 (5'-GAGCAAGGTGAGATGACAGGAG-3') and 146RP 1718 (5'-CAGTGACAACGTCGAGCACAGC-3') in a nested PCR.

Bioinformatics analyses. The Genome Cryptographer software (http://shark.ucsf.edu/gc/) was used to identify and graphically visualise the distribution of transposable elements and repeated regions, and results were checked with Repeatmasker2 (http://www.repeatmasker.org/) screens with maximum sensitivity level. Information on TEs was also found on http://www.girinst.org/index.html.

MAR-Wiz (http://www.futuresoft.org/MAR-Wiz/), SMARTest (http://www.genomatix.de/products/SMARTest/index.html), Marscan and Fuzznuc (EMBOSS, http://emboss.sourceforge.net/) were used to predict putative MARs in the flanking genomic DNA. SMARTest and MAR-Wiz show the MARs as defined regions (both analyses performed at default settings). Marscan searches for a bipartite MAR recognition signature (MRS) consisting of two individual sequences of 8 bp (AATAAYAA)

and 16 bp (AWWRTAANNWWGNNNC) in various configurations (default setting). Fuzznuc identifies a degenerate 21-bp MRS (TAWAWWWNNAWWRTAANNWWG) (both strands were searched, two mismatches allowed, but motive has to end with G). The last 12 nucleotides of the 21-bp MRS are a part of the 16-bp MRS. When the 16-bp MRS was present within a MAR predicted by another program (SMARTest or MAR-Wiz) or at the same position (Fuzznuc), it was considered to be an overlapping MAR. MARs identified by SMARTest and MAR-Wiz were scored as overlapping if they partly or entirely covered the same region.

RT-PCR for detection of putative antisense/promoter transcripts. Total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen). First-strand cDNA synthesis was performed on 1 µg of deoxyribonuclease-I-treated RNA with 5 pmol of a specific primer, 5 nmol dNTPs (each), 1 X AMV reverse transcription (RT) Reaction Buffer l, 5 U AMV reverse transcriptase and 1 µl recombinant ribonuclease inhibitor (20–40 U/µl) at 48 °C for 45 min. A parallel reaction was prepared without the RT enzyme. 1.0 µl (1/25 vol) of each reaction mixture was used for PCR. 3'-end primers in the PCR reactions were nested on the primers used for first-strand synthesis. The following primers were used for detection of aberrant promoter transcript; ab5' (5'-TGAGCGGAGAATTAAGGGAGTC-3'); ab3'-1 (5'-TGTTGTGCCCAGTCATAGCC-3'); ab3'-2 (5'-AACCTGCGTGCAATCCATC-3'); for detection of antisense *nptII* transcript: as5' (5'-TCATAGGCGTCTCGCATATCTC-3'); as3'-1 (5'-CTGCTTGCCGAATATCATGG-3'); as3'-2 (5'-CGCTTCCTCGTGCTTTACG-3'); and for the actin control; act2int3_sense (5'-TCAGGAAGGATCTCTATGGAAAC-3'); act2int3_antisense (5'-TTCCTGTGAACAATCGATGG-3').

Analysis of DNA methylation by bisulphite sequencing. The following protocol modified from previous studies [41–43] was used: 1 µg DNA was digested with PstI and EcoRI, phenol/chloroform extracted, ethanol precipitated, dissolved in 20 µl H_2O and incubated with 1/9 vol of 3 M NaOH for 20 min at 37 °C. Then, 208 µl urea/bisulphite solution (6.24 M urea/2 M metabisulphite) and 12 µl 10 mM hydroquinone were added and incubated for 18 h at 55 ºC with a 30-s denaturation step at 95 ºC every 2 h. Resulting DNA samples were desalted, eluted in 20 µl H_2O and incubated for a second time with 1/9 vol of 3 M NaOH for 20 min at 37 ºC. The reaction was neutralised by adding NH_4OAc to a final concentration of 3 M, ethanol precipitated and resuspended in 50 µl H_2O. Bisulphite-treated DNA (1–5 µl) was used in a subsequent PCR reaction with the degenerate primers 5'GS (5'-YATGAGYGGAGAATTAAGGGAGT-3') and 3'GS (5'-CCRAATARCCTCTCCACCCAA-3') as designed by Aufsatz et al. [44]. A minimum of ten clones of each PCR product were sequenced.

Generation of transgenic *crypt::GUS* plants. The *pPZP P4 crypt*::*GUS* and *pPZP P10 crypt*::*GUS* constructs were made using Gateway cloning technology (Invitrogen). For line P4, an 865-bp fragment was amplified from *Arabidopsis* ecotype C24 DNA using the Gateway *att*-modified primers P4 crypt F (5'-*att*B1-ATCGTCTGCTTTACCACTCTCCTC-3') and P4 crypt R (5'-*att*B2-TGAACTATATCTGAATCTTATTGATCG-3'). For line P10, a 484-bp fragment was amplified from the plasmid pPCV002 [45] using the *att*-modified primers P10 crypt F (5'-*att*B1-ATGAGTATGATGGTCAATATGGAG-3') and P10 crypt R (5'-*att*B2-CCGACAGAGGTGTGATGTTAG-3'). Amplified fragments were introduced into the pDONR/Zeo Gateway entry vector and recombined into the pPZP211GGAWI destination vector [38], generating the constructs *pPZP P4 crypt*::*GUS* and *pPZP P10 crypt*::*GUS*. These were transferred to the *Agrobacterium tumefaciens* strain C58C1 pGV2260 and used to transform *Arabidopsis* (ecotype C24) by the floral dip method [46]. Transformants were selected by germinating seeds on plates containing 50 µg/ml Km.

Histochemical GUS assay. The GUS assay was performed after a modified protocol from Grini et al. [47]. Plant material was incubated in staining buffer (50 mM $NaPO_4$, pH 7.2, 2 mM potassium ferrocyanide, 2 mM potassium ferricyanide, 0.1% Triton X-100, 10 mM X-Gluc) in the dark at 37 °C for 48 h, treated with 1:2 CH_3COOH:EtOH for 10 min and washed with 70% EtOH and then sterile H_2O before mounting on microscope slides in a clearing solution of 8:2:1 chloral hydrate:water:glycerol. Slides were inspected with a Zeiss Axioplan2 imaging microscope equipped with differential interference contrast optics and a cooled Axiocam camera imaging system.

Results

Neither transcriptional interference from the *GUS* gene or neighbouring genes, nor presence of vector backbone could explain the silencing found. In the present work, we first focused on the genomic surroundings of T-DNA insertions in five lines displaying silencing and three without silencing (table 1). In the K lines, the *nptII* gene is found close to the T-DNA left border (LB) and positioned downstream of and in the same orientation as the *GUS* gene also present in the T-DNA (fig. 1A). In the P lines, the *nptII* gene is found close to the T-DNA right border (RB) and is oriented in opposite direction to the *GUS* gene (fig. 1B, C). Lines *ida*,

Table 1. Position and orientation of T-DNA insertions relative to upstream (Ups) and downstream (Dws) neighbouring genes.

Line	Closest Ups		Closest Dws	
	Gene	cDNA ac. no.	Gene	cDNA ac. no.
P4	At1g71710(C) –6470 tl	AY048296	At1g71720 (W) –461tl	AI997850 EST[a, b]
P10	At1g23980(C) –138tl	AK119101	At1g23990 (W) –4532tl	n.d.[c]
K11	At3g48690(C) +56tl	AY064980	At3g48700 (C) +722st	AK118967
K14	At1g60080(C) –1101tl	NM_ 104700	At1g60090 (W) –142tl	n.d.[c]
K15	At3g18070(W) +25st	n.d.[c]	At3g18080 (W) –1528tl	AF360240
ida	At1g68780(W) +1026st	AF361616	At1g68765 (W) –392tl	AY087883
747	At4g36550(C) –2160tl	F14300 EST/Y	At4g36560 (W) –2763tl	n.d.[c]
755	At5g05100(C) –115tl	AY059918	At5g05110 (C) +560st	AF370168

[a] No cDNA, accession number given for EST.
[b] No cDNA, expression confirmed by Yamada et al. [48].
[c] Expression has not been detected.

Gene annotations are according to the Munich Information Centre for Protein Sequences (MIPS) (http://mips.gsf.de/), and the orientation of the genes indicated by Watson (W, sense) and Crick (C, antisense). The distance to the T-DNA insertion point is given in number of base pairs upstream (–) or downstream (+) of the translations start (tl) or stop (st) codon of the neigbouring genes.

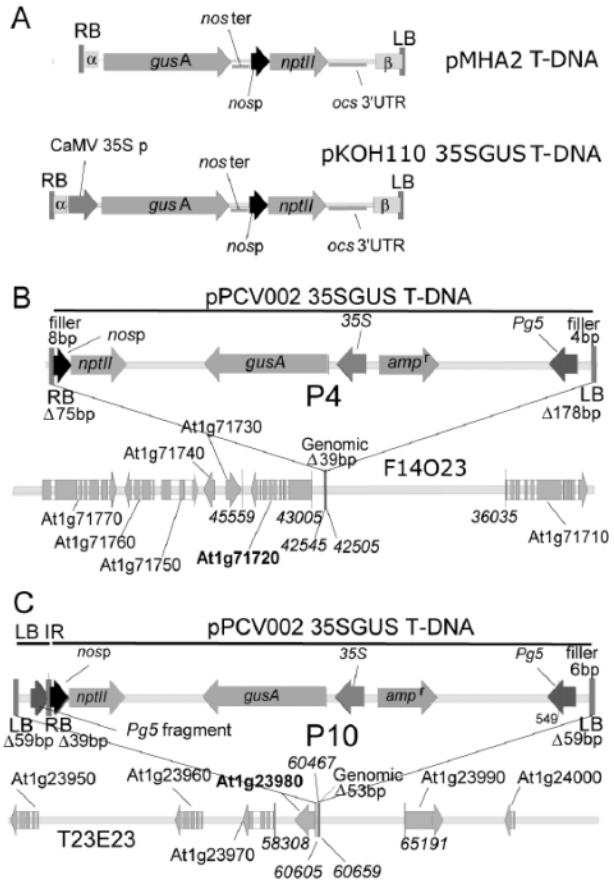

Figure 1. Schematic representation of the T-DNA regions and insertions. LB and RB, left and right T-DNA border; α and β, T-DNA sequences; CaMV 35S and *nos*p, Cauliflower Mosaic Virus 35S promoter and *nos* promoter, respectively; *nos* ter and *ocs* 3'UTR, terminators and polyadenylation signals; *gus*A, *nptII* and *amp*^{r}: β-glucuronidase A, *neomycin phosphotransferase* (Km resistance) and ampicillin resistance gene, respectively; Pg*5*, truncated promoter of T_L-DNA gene *5*. (*A*) T-DNAs of the pMHA2 and pKOH110 35SGUS vectors used to transform promoter-trap lines and K lines, respectively. (*B*) Position and characteristics of T-DNA insertion in line P4. The pPCV002 35SGUS T-DNA in the upper part is shown with deletions (△) in the T-DNA and filler sequences at the junctions between plant and T-DNA. The lower part depicts a part of the BAC clone F14O23 (accession no. AC012654) surrounding the T-DNA, with a deletion at the insertion site. (*C*) Position and characteristics of T-DNA insertion in line P10. The pPCV002 35SGUS T-DNA, with a 490-bp inverted repeat of the LB side (LB IR) outside the RB side, is shown in detail in the upper part. The lower part depicts a part of the BAC clone T23E23 (accession no. AC002423), with a deletion in the insertion site. The nearest genes with their exons are indicated with boxes/arrows (indicating transcriptional direction). Relevant BAC positions are given in italics.

SENAPE4/775 and LIGUINE1/747 are without *nptII* silencing, and have a promoterless *GUS* gene, but were included as they have other features in common with our silenced lines (fig. 1A and below).

The positions of the T-DNA insertions relative to annotated genes are summarized in table 1. The T-DNAs of K14 and K15 were inserted just upstream and downstream, respectively, of the coding sequence (CDS) of genes for which no expression data are available. The T-DNAs of P4, P10 and *ida* were found integrated in promoter regions. In K11 and SENAPE4/775, the T-DNAs were inserted in transcribed regions [CDS and the 5' untranslated region (UTR), respectively], causing *GUS* reporter gene expression in the SENAPE4/775 line [39, 40]. The *GUS* gene was also expressed in the LIGUINE1/747 line, although the T-DNA was found in an intergenic region (IGR), suggesting insertion in an undiscovered gene or next to a cryptic promoter [39, 40]. No silencing was found in the SENAPE4/775 and LIGUINE1/747 lines, arguing that transcriptional interference from the *GUS* gene is unlikely to explain *nptII* silencing in the K lines. The third non-silenced line, *ida*, is comparable to K11 in that long stretches of BVB were found continuous from the LB of the T-DNA, in addition to an intact T-DNA copy [14, 38]. Thus, the presence of BVB is in itself not sufficient to induce silencing [14].

In most cases, transcription from both of the genes flanking the T-DNA insertion could be verified either by a cDNA record, ESTs or by whole genome array (WGA) expression analysis [48] (table 1). However, the orientation of the *nptII* gene relative to the flanking genes was not such that it allowed generation of *nptII* antisense transcripts by read-through transcription from neighbouring genes, except for K11. Antisense expression of the gene tagged in K11 (documented by MPSS [49] and by WGA [48]) could alternatively generate an antisense *nptII* transcript in this line. However, in both cases, transcription would then have had to proceed through the more than 5-kb-long BVB.

No single, common factor in the genomic surrounding of the T-DNAs could be suggested to explain the silencing in the investigated lines. Genomic regions of ±20 kb flanking the transgenes were analysed by bioinformatics tools. While T-DNA and the BVB contain more than 55% GC, the GC content of the insertion sites both in silenced and non-silenced lines was close to the genomic average, i.e. 35% [50] (table 2). A lower GC fraction in a region 50–100 bp just around the integration sites in lines P10 and K14 may reflect insertion in promoter regions. Conversely, in K11, T-DNA is inserted in an exon of a particularly GC-rich gene. In general, CDSs have a higher GC content (44%) [50].

Four different programs developed for prediction of putative MARs were used [51–55]. Only overlapping MARs,

Table 2. Summary of identified features in non-silenced and silenced lines.

Line	Position	Extra DNA[a]	GC[b]	MARs[c]	TE[d]
P4	P	–	m.r.	1/0.1	+
P10	P	inverted T-DNA	<30	0/0.1	+
K11	exon	BVB	>50	0/0.15	+
K14	P/5'UTR	–	<30	0/0.2	+
K15	3'UTR?	–	m.r.	2/0.1	–
ida	P	BVB	m.r.	2/0.25	+
747	IGR	–	m.r.	1/0.25	–
755	5'UTR	–	m.r.	2/0.23	–

[a] truncated T-DNA or BVB in addition to intact T-DNA.
[b] GC composition ± 50 bp around the integration sites. m.r., mid range (between 30 and 40%).
[c] putative overlapping MARs within ± 2 kb of the integration site and number of overlapping MARs per kb in the analysed ± 20 kb.
[d] presence (+) or absence (–) of transposable elements in the analysed ± 20 kb.
P, promoter; IGR, intergenic region.

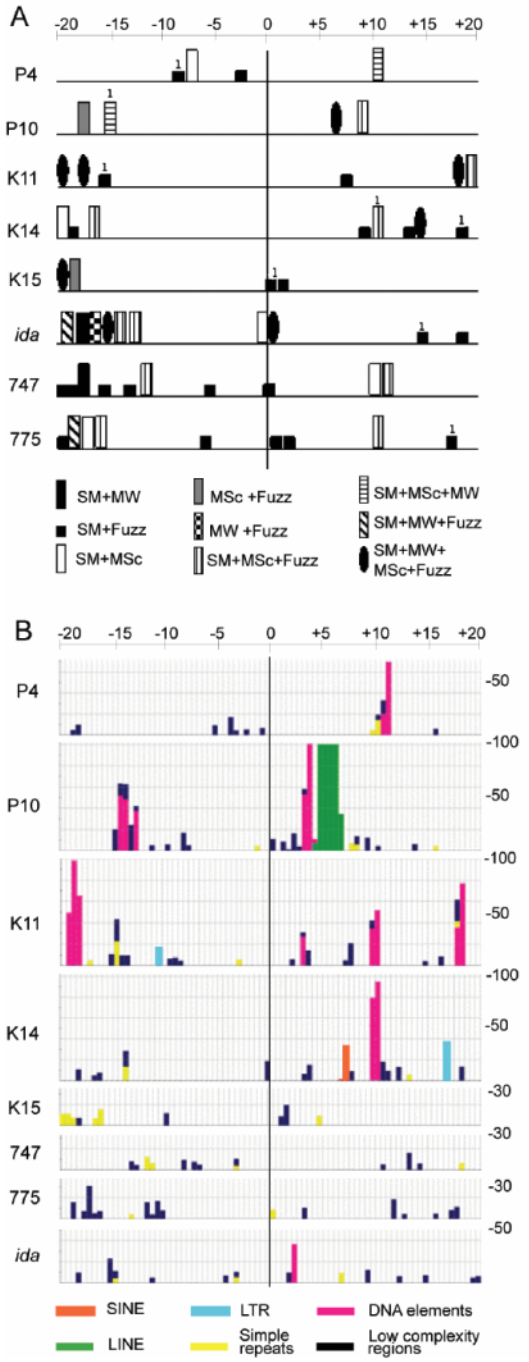

Figure 2. Putative MARs and repeated elements in plant DNA flanking T-DNA insertions. The transgenes are integrated at position 0. (*A*) Distribution of putative MARs identified in flanking plant DNA (± 20 kb). Overlapping MARs predicted at the same location by at least two programs, as indicated, are shown. Fuzznuc (Fuzz) signatures with a single mismatch compared to the consensus are indicated with 1, while the other Fuzznuc signatures have two mismatches. SM, SMARTest; MW, MAR-Wiz; MSc, Marscan (see text and Materials and methods). (*B*) Schematic overview of repeated elements in flanking plant DNA (± 20 kb) generated with Genome Cryptographer software at 500-bp resolution. The Y-axis designates the percentage of a 500-bp DNA segment an element occupies. Simple repeats and low complexity regions typically consist of various repeated consensus sequences of 2–4 nucleotides, or AT-rich regions of 20–50 nucleotides. See text for details on transposable elements.

usually consisting of a MAR predicted by SMARTest or MAR-Wiz, in combination with one or more short MRSs found by Marscan and Fuzznuc (fig. 2A), are reported here. About twice as many putative MARs were found in the investigated window in the three non-silenced lines compared to the silenced lines (table 2). These MARs were also found closer to the insertion site in the non-silenced lines.

Various simple repeats, low-complexity regions and transposable elements were identified with Genome cryptographer and Repeatmasker2 software programs (fig. 2B). Low-complexity regions and simple repeats were found both in silenced and in non-silenced lines. A notable difference between silenced and non-silenced lines was the occurrence of transposable elements (non-autonomous DNA transposons, LINEs, SINEs and remnants of retroelements) around the insertion site.

One may speculate that a long distance to MARs, in combination with closeness to transposable elements, might influence transgene expression in our particular silenced lines. But, obviously, general conclusions cannot be drawn from the *in silico* analysis on our restricted material and lines with different constructs, and a more detailed molecular investigation to elucidate the silencing phenomena in our lines was undertaken.

A transcript covering the *nos* promoter was detected in lines P4 and P10. In our lines displaying silencing, the frequency of Km-sensitive, i.e. *nptII*-silenced seedlings was low and varied between sibling sublines and generations [14, 36]. Only in two of these lines was a molecular comparison between Km-sensitive and -resistant plants possible. In line P4, 100% of the seedlings of one sibling subline homozygous for the T-DNA insertion consistently showed type III silencing when grown on Km plates, i.e. plantlets had green-and-white spotted leaves that were often pointed and deformed, indicating *nptII* silencing in a fraction of the cells only [36]. In the T_5 generation of line P10, one homozygous sibling subline showed more than 80% type I silencing, i.e. seedlings had a phenotype identical to wild type plants grown on Km.

The T-DNAs of P4 and P10 are positioned just upstream of the start codon of genes At1g71720 and At1g23980 (fig. 1B, C). Cloning of the DNA adjacent to the right border of the P10 T-DNA revealed the presence of a 490-bp left border T-DNA fragment in inverted orientation (fig. 1C). This fragment contains a part of a truncated T_L-DNA Pg5 promoter [45].

In line P10, the structure and position of the T-DNA insertion led to two hypotheses regarding potential silencing-triggering RNA species. First, transcription directed by the putative interrupted endogenous At1g23980 promoter could possibly proceed through the entire T-DNA from its left border side, generating antisense transcripts of the *nptII* gene. cDNAs with a 5' UTR of only 14 bp (AK119101.1, RAFL21-45-G23) have been identified for At1g23980. Second, transcription driven by the Pg5 promoter fragment in the left border inverted repeat could possibly generate aberrant *nos* promoter (*nos*p) transcripts.

These hypotheses were tested by RT-PCR (fig. 3A) on total RNA from the silenced P10 subline germinated on medium without Km, and from control groups consisting of Km-resistant plants from sibling sublines of P10 and P4 without silencing. No antisense RNA was detected

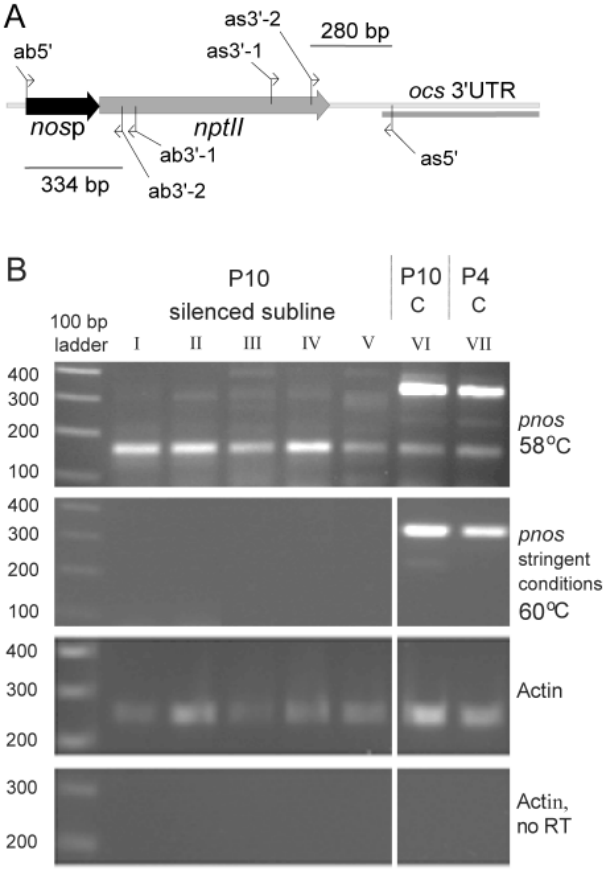

Figure 3. Detection of aberrant transcripts. (*A*) Primers for RT reaction and PCR at the RB side of pPCV002 35SGUS T-DNA to detect putative *nos* promoter transcripts (RT reaction: ab3'-1; PCR: ab3'-2 + ab5') and antisense RNA (RT reaction: as3'-1; PCR: as3'-2 + as5'). Expected PCR product sizes are indicated (horizontal lines). (*B*) RT-PCR results with ab primers under non-stringent (annealing temperature 58 °C) and stringent (60 °C) conditions. Lanes I–V, P10 single plants; lane VI, P10 non-silenced control; lane VII, P4 non-silenced control. Expected product sizes are 334 bp for the *nos*p transcript, and 255 bp for the actin mRNA control. The 180-bp bands in the upper panel represent an unspecific amplification of a ribosomal chloroplast cDNA at non-stringent conditions. A negative control without the RT enzyme was run for all reactions, but is only shown for actin. RT-PCR aimed at detecting antisense transcripts was negative (data not shown).

(results not shown). However, the presence of a sense *nos*p transcript (334 bp) was clearly demonstrated (fig. 3B). Under stringent conditions (annealing temperature 60 °C), the *nos*p transcript was detected, contrary to our expectations, only in the Km-resistant controls. Under less stringent conditions (annealing temperature 58 °C), weak PCR products of expected size were amplified also from the five plants of the silenced subline of P10. Sequencing verified the *nos*p identity of these bands.

The *nos* promoter is methylated in silenced plants. The DNA methylation status of the same region as that covered by the RT-PCR product (fig. 4A) was investigated using bisulphite-mediated genomic DNA sequencing. In the P4 line, seedlings from two different siblings (s-1 and s-2) of the subline showing ~100% silencing were germinated without Km and compared to Km-resistant seedlings from sublines (A and B) without silencing (controls). While most of the P4 control sequences showed little or no cytosine methylation, the number of methylated cytosines from the silenced P4 sublines samples ranged from 3 to 38, with an average of 18 methylated residues (fig. 4B).

Substantial variation in the methylation pattern between sequences is evident in the samples from line P4, as has been observed in earlier reports from the same region [44, 56, 57]. This could be due to a differential methylation pattern in different seedlings or in our case in different cells, since the P4 subline shows a type III (spotted) silencing pattern. In line P10, where one subline showed a high level of type I (uniform) silencing, we investigated single plants. Compared to line P4, a generally higher methylation level was observed, with strikingly uniform methylation patterns, i. e. 10 out of 12 sequences from the same plant had the same 35 residues methylated, while the last two clones showed a slightly different pattern and a lower methylation level (22 residues) (fig. 4C). The P10 Km-resistant control (multiple plants, as for line P4) showed a significantly lower methylation level, although this level was considerably higher than in the control of line P4.

Both for P4 and P10, there was a higher overall level of methylation in symmetrical compared to asymmetrical cytosine positions, particularly in CpG positions (fig. 4B, C). Methylation levels were pronouncedly higher in the *nos*p region than in the transcribed region of *nptII*. This difference was less evident for methylation in cytosines in a CpG context, compared to methylation in CpNpG and CpNpN contexts.

DNA sequences adjacent to *nos*p in lines P10 and P4 can drive reporter gene expression. To explain the existence of the *nos*p transcripts, we searched for promoter elements in the DNA sequences upstream of *nos*p (fig. 5). Putative core promoter elements were found in both

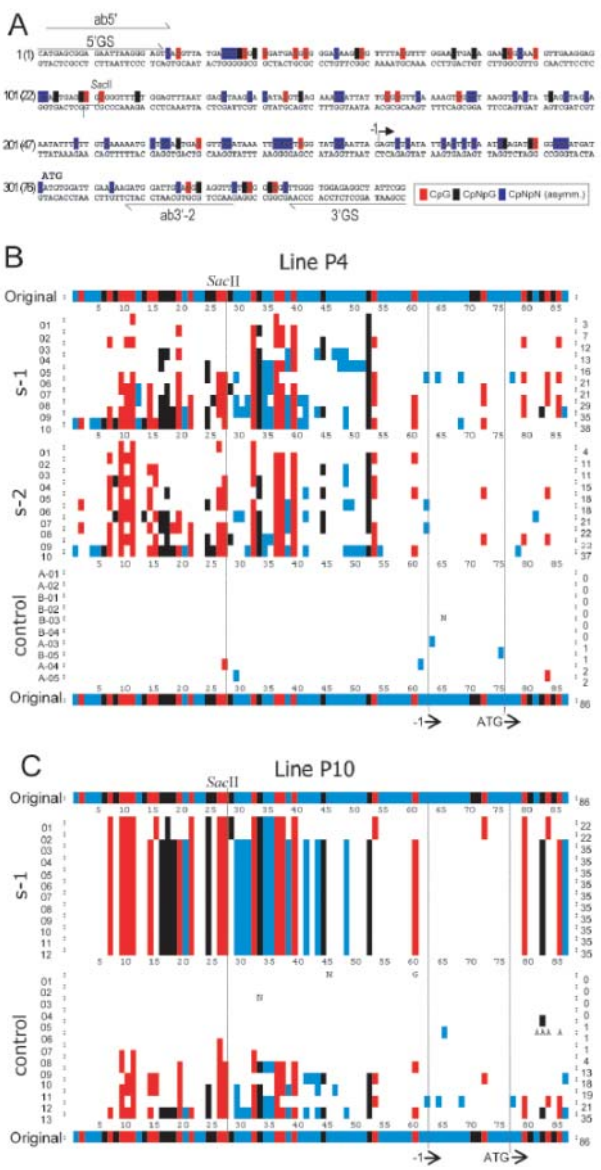

Figure 4. DNA methylation analysis by bisulphite sequencing. (*A*) Original (unconverted) sequence, consisting of the *nos* promoter and a part of the *nptII* gene. Numbers to the left are positions in the complete sequence and in parentheses the corresponding positions of the upper-strand cytosines represented in *B*. The sequence of the primers 5'GS and 3'GS used for sequencing and (for comparison) ab5' and ab3'-2 used for the RT-PCR (fig. 3A), the restriction site of *Sac*II, the transcription start site (bent arrow at –1) and the *nptII* start codon are indicated. The cytosines in the upper strand of the sequence are marked in colour codes as indicated. (*B*) Bisulphite sequencing results for P4. Sibling lines with silencing (s-1 and s-2), and controls A and B (no silencing). (*C*) Bisulphite sequencing results for P10. Single plant with silencing (s-1), and control (multiple plants, no silencing). (*B, C*) Only the 5'-methylcytosines in the upper strand of the *nos*p-*nptII* sequence are given (except for the Original sequence, where all cytosines are given), and are marked in colour codes corresponding to *A*. Total numbers of methylated residues are given to the right of each sequence. The digestion site of *Sac*II, the transcription start site (–1) and the ATG start codon are indicated.

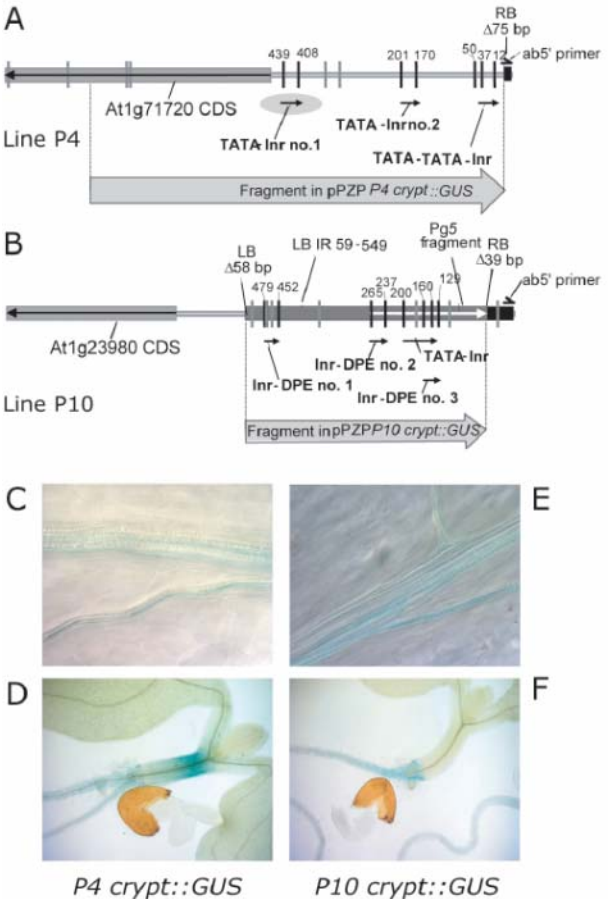

Figure 5. Cryptic promoter analyses. (*A, B*) The genomic regions 1 kb upstream of the RT-PCR ab5' primer in lines P4 and P10. The sequences included in the *pPZP P4/P10 crypt::GUS* constructs are shown with grey arrows. Medium-dark-grey boxes indicate CDSs of endogenous genes, with arrows indicating direction of transcription; dark-grey (LB inverted repeat in line P10) and black boxes (RB side) indicate T-DNA. The truncated Pg5 promoter in line P10 is marked with a white arrow. Possible composite cryptic core promoters (black arrows) with combinations of TATA elements, transcription initiator elements (Inr), or downstream promoter elements (DPE) (marked as vertical black lines) are shown [69], as are CAAT boxes (grey lines). Numbers indicate distance in bp relative to the ab5' primer. The only composite cryptic promoter with a possibility for transcription overlapping with transcripts from any of the putative endogenous promoter elements of the At1g71720 gene is marked with a grey circle. (*C–F*) GUS expression from the constructs *pPZP P4 crypt::GUS* (*C*, *D*) and *pPZP P10 crypt:GUS* (*E*, *F*). (*C, E*) Expression in leaves of 7-week-old primary transformants (x40 magnification). (*D*, *F*) Expression in 3-week-old seedlings of the next generation (x5 magnification).

lines, in P4 in the genomic DNA, and in P10 in the short left border repeat. Fragments containing these putative cryptic promoters were amplified by PCR and inserted upstream of the *GUS* reporter gene in the vector pPZP211G-GAWI (fig. 5A, B). The resultant constructs *pPZP P4 crypt*::*GUS* and *pPZP P10 crypt*::*GUS* were used to generate transgenic *Arabidopsis* plants, and Km-selected transformants were inspected for GUS expression. For both constructs, weak expression was observed in vascular tissue in the lower parts of the leaves of 4 out of 13 investigated primary transformants (fig. 5C, E). In seedlings of the next generation, expression was consistently confined to the roots of plants with the P10 construct, but was also observed in the hypocotyl of plants with the P4 construct (fig. 5D, F). Both of these expression patterns were observed in several independent transformants/sublines, clearly demonstrating that the cloned fragments were able to drive transcription.

Discussion

The *nos* promoter may be transcribed from cryptic promoter elements in P4 and P10. Our *in silico* analyses did not reveal any single triggering factor common to all five lines with silencing, consistent with findings by others [10, 15]. Therefore, it is tempting to suggest that a combination of line-specific factors, e. g. those summarized in table 2 and/or others yet unidentified, are contributing to silencing. However, since the *in silico* analysis was inconclusive, molecular investigations of putative position effects were undertaken for the P lines for which sublines with a complete or very high degree of silencing were at hand. The *nptII* gene in the P lines is positioned differently than in the K and pMHA2 lines, with the *nos*p pointing inwards from the right border (fig. 1). The T-DNAs in line P4 and line P10 are positioned in the promoter regions of transcriptionally active genes, with the promoters pointing in the opposite direction to *nos*p-*nptII*. Using RT-PCR, unusual transcripts through the *nos*p were detected in these two lines (fig. 3).

Aberrant transcripts have been implicated in many models of silencing [see e. g. refs. 58–60]. An *in silico* search for core promoter elements suggested that the observed transcripts were generated by transcription from adjacent cryptic promoters present in the upstream region of At1g71720 in line P4, and in line P10, the inverted T-DNA fragment containing a truncated Pg5 promoter (fig. 5A, B). The *P4* and *P10 crypt::GUS* transgenics expressing GUS demonstrate that these DNA regions are able to drive expression of adjacent sequences, and thus provide an explanation for the presence of the *nos*p transcripts.

The methylation pattern in the *nos* promoter sequence indicates an RNA-mediated mechanism. In line P4 and P10, we could compare methylation levels of the region covered by the detected aberrant *nos* transcript in sublines without silencing and siblings with a high degree of silencing. In line P10, but not P4, an association between methylation in the *Sac*II site and silencing is evident. These results are consistent with previous restriction endonuclease analysis on our material [14, 36,

37], but contradict studies showing a consistent association [e. g. ref. 61]. This may partly be explained by the different silencing phenotypes observed in the different lines (type I in P10 and type III in P4 [36]), assuming that only the heavily methylated clones represent silenced tissue. However, following this line of reasoning, one might expect a type III silencing phenotype in the control sample of P10, where a substantial level of methylation was also seen. On the contrary, these plants appeared completely Km resistant. This suggests that the significance of DNA methylation in the two lines is different, i. e. a certain level that is sufficient or indicative of the type III silencing observed in line P4 is not sufficient or indicative for the type I silencing observed in line P10. Possibly, such differences might be connected to the particular genomic environment of each T-DNA (e. g. factors summarized in table 2), or the extra inverted T-DNA fragment in line P10.

The highest incidence of methylation in the silenced sublines of P10 and P4 was in the *nos*p region (fig. 4B, C), suggesting that the observed Km sensitivity is due to TGS of the *nptII* gene [44, 56, 62].

Methylation in line P10 and P4 frequently appears in the same symmetrical cytosine residues (CpG and CpNpG) within and across different sublines of the same transformant (fig. 4B, C), indicating maintenance of methylation patterns established in one of the earlier generations. More striking, however, are the highly uniform patterns of methylation in single plants (P10) also in asymmetrical cytosines. This suggests that once a methylation pattern is established in any particular cell, it is faithfully inherited by all mitotic descendants, partially independent of mechanisms reliant on symmetry between complementary DNA strands.

Such independence of symmetry can be explained by the repeated activity of de novo DNA methyltransferases, guided to target sequences by RNA in each replicative cycle [44, 56, 63, 64]. The presence of unusual *nos*p transcripts and *nos*p methylation patterns similar to that reported for RNA-induced silencing of *nos*p is highly suggestive of an RNA-based silencing mechanism in P10 and P4, or, more specifically, RNA-directed DNA methylation (RdDM) [44, 56, 57, 65].

Stochastic triggering and inheritance of silencing may be explained by a mechanism where transcription of the *nos* promoter is the initiating factor. The establishment of silencing is of a stochastic nature in our lines. As the *nos*p transcript actually was detected at higher levels in non-silenced siblings (fig. 3B), a sense transcript of a promoter sequence is not in itself sufficient for silencing, consistent with findings of others [66]. However, in our study, the detected *nos*p transcripts are transcribed from the very region that becomes methylated. These *nos*p transcripts may be truncated or otherwise aberrant, causing recognition by an RdRP, synthesis of dsRNA and post-transcriptional degradation (see suggested model in fig. 6). In the next step, either the dsRNA or derived siRNAs could direct DNA methylation of homologous regions (fig. 6). An initial level of aberrant transcripts may induce a low level of methylation that could recruit chromatin factors, occasionally escalating into a situation where *nos*p is no longer active. Importantly, aberrant transcripts are likely to be processed and/or degraded following the RdDM process (fig. 6), explaining the lower level of *nos*p transcripts in silenced plants (fig. 3B). Continued transcription of *nos*p must, however, be suggested to explain the consistent presence of asymmetric methylation through mitosis and/or meiosis.

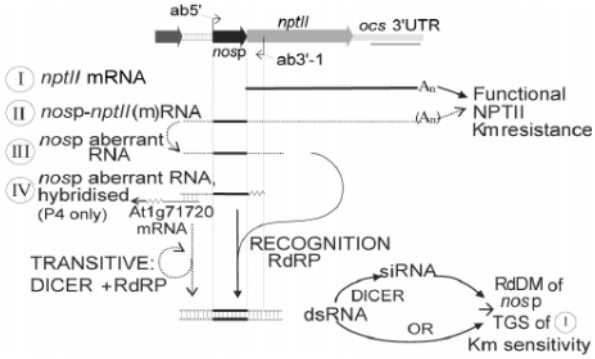

Figure 6. Model for RNA-mediated silencing in lines P4 and P10. Top of figure corresponds to figure 3B with a suggested cryptic promoter represented by a dark-grey arrow. Transcript I is the normal *nptII* transcript; putative transcripts II, III and IV may be detected by the RT-PCR primers used. Uncertain points of origin and termination are indicated with stippled segments. A transcript II, originating upstream of the *nos* promoter, giving a full-length *nptII* transcript with a longer 5'UTR than normal, may possibly be polyadenylated and functional; transcript III is a truncated and/or otherwise aberrant version of II; transcript IV (in line P4 only) may be either full-length or truncated in the 3'-end, but hybridised in the 5'-end with the mRNA of At1g71720. Both transcript III and IV may be recognised as aberrant by an unprimed RdRP, which synthesises dsRNA. Transcript IV may enter a separate pathway that involves initial targeting of the hybridised region, but that extends into the *nos*p region by a transitive silencing mechanism. dsRNA synthesis is concentrated in the region in and possibly upstream of the *nos* promoter, but may also extend into the *nptII* coding region, probably depending on the length of the aberrant RNA. Either long dsRNA or siRNAs (dsRNA cleaved by Dicer-like enzymes) may be involved in the next steps, causing DNA methylation in the *nos* promoter region. This is likely to result in TGS of transcript I, either directly, or through subsequent reinforcement by interaction with chromatin components and increased levels of methylation in symmetrical cytosines. The RdDM process is likely to be ongoing also after shut-down of the *nos* promoter, maintaining or reinforcing silencing.

Construct structure may influence transgene susceptibility to position effects. Other groups [see e. g. ref. 4] have first selected single-copy lines, and then looked for silencing of the *nos*p-*nptII* cassette without identifying a

single case of silencing of the *nptII* gene. Our lines were selected by a different approach: we first identified silenced lines, and thereafter the number of T-DNA copies [14, 36]. In addition, Schubert et al. [4] used a construct where the *nos* promoter is more distant from the T-DNA border than is the case for the construct in our P lines. Promoter-trap experiments clearly demonstrate that transcription can occur from the adjacent DNA into a T-DNA. This is exemplified for the promoterless *GUS* gene in SENAPE4/775, most likely transcribed from a natural promoter, and in LIGUINE1/747, most likely from a cryptic promoter [39]. In contrast to the two P lines, the frequency of silencing was low in all sublines and generations tested for the K lines [14, 36]. This might indicate different silencing mechanisms. Silencing in our K lines is unlikely to be due to the generation of aberrant *nos*p transcripts, since the *nos* promoter is positioned in the middle of the T-DNA in these lines. However, there are interesting data from other studies using constructs similar to our P construct. A thorough investigation of the effects of reporter gene orientation and position on transgene expression levels (30–60 plants for each construct) has previously demonstrated that positioning of the *nos* promoter close to the right T-DNA border resulted in higher intertransformant variability and a higher frequency of transformants with very low expression [67]. This is consistent with our model for position-dependent silencing due to the increased risk of production of aberrant promoter transcripts when the promoter of the transgene is close to the border of the inserted DNA. We suggest that this mechanism may also be at play in a recent study revealing a high occurrence of silenced T-DNA integrations in transformants identified by PCR screening instead of *nptII* selection [68]. In this work, a vector with the *nos*p-*nptII* gene located at the right border was used. The lines identified in these studies would therefore be suitable for testing the general significance of our model.

Acknowledgements. A grant to I. S. Mercy was provided by a Strategic University Program funded by the Research Council of Norway (project number 129525/420). We thank T. J. Meza for optimisation of the bisulphite sequencing protocol, and M. Skårn for valuable discussions regarding the use of bioinformatics programs to analyse features of the plant DNA flanking T-DNA insertions. We thank S. H. Engebretsen, R. Falleth and K. E. Rakkestad for technical assistance. This work has been facilitated by the Norwegian Arabidopsis Research Centre, a national technology platform supported by the Research Council of Norway.

1 Vaucheret H. and Fagard M. (2001) Transcriptional gene silencing in plants: targets, inducers and regulators. Trends Genet. **17:** 29–35
2 Matzke M. A., Aufsatz W., Kanno T., Mette M. F. and Matzke A. J. (2002) Homology-dependent gene silencing and host defence in plants. Adv. Genet. **46:** 235–275
3 Fagard M. and Vaucheret H. (2000) (Trans)gene silencing in plants: how many mechanisms? Annu. Rev. Plant Physiol. Plant Mol. Biol. **51:** 167–194
4 Schubert D., Lechtenberg B., Forsbach A., Gils M., Bahadur S. and Schmidt R. (2004) Silencing in *Arabidopsis* T-DNA transformants: the predominant role of a gene-specific RNA sensing mechanism versus position effects. Plant Cell **16:** 2561–2572
5 Chuang C. F. and Meyerowitz E. M. (2000) Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. Proc. Natl.Acad. Sci. USA **97:** 4985–4990
6 Waterhouse P. M., Graham M. W. and Wang M. B. (1998) Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc. Natl.Acad. Sci. USA **95:** 13959–13964
7 Stam M., Viterbo A., Mol J. N. and Kooter J. M. (1998) Position-dependent methylation and transcriptional silencing of transgenes in inverted T-DNA repeats: implications for post-transcriptional silencing of homologous host genes in plants. Mol. Cell. Biol. **18:** 6165–6177
8 Vaucheret H., Beclin C. and Fagard M. (2001) Post-transcriptional gene silencing in plants. J. Cell Sci. **114:** 3083–3091
9 Matzke M. A., Matzke A. J., Pruss G. J. and Vance V. B. (2001) RNA-based silencing strategies in plants. Curr. Opin. Genet. Dev. **11:** 221–227
10 De Buck S., Windels P., De Loose M. and Depicker A. (2004) Single-copy T-DNAs integrated at different positions in the *Arabidopsis* genome display uniform and comparable beta-glucuronidase accumulation levels. Cell. Mol. Life Sci. **61:** 2632–2645
11 Day C. D., Lee E., Kobayashi J., Holappa L. D., Albert H. and Ow D. W. (2000) Transgene integration into the same chromosome location can produce alleles that express at a predictable level, or alleles that are differentially silenced. Genes Dev. 14: 2869–2880
12 Elmayan T. and Vaucheret H. (1996) Expression of single copies of a strongly expressed 35S transgene can be silenced post-transcriptionally. Plant J. **9:** 787–797
13 Meyer P., Linn F., Heidmann I., Meyer H., Niedenhof I. and Saedler H. (1992) Endogenous and environmental factors influence 35S promoter methylation of a maize A1 gene construct in transgenic petunia and its colour phenotype. Mol. Gen. Genet. **231:** 345–52
14 Meza T. J., Stangeland B., Mercy I. S., Skarn M., Nymoen D. A., Berg A. et al. (2002) Analyses of single-copy *Arabidopsis* T-DNA-transformed lines show that the presence of vector backbone sequences, short inverted repeats and DNA methylation is not sufficient or necessary for the induction of transgene silencing. Nucleic Acids Res. **30:** 4556–4566
15 Qin H., Dong Y. and Arnim A. G. von (2003) Epigenetic interactions between *Arabidopsis* transgenes: characterization in light of transgene integration sites. Plant Mol. Biol. **52:** 217–231
16 Schotta G., Ebert A., Dorn R. and Reuter G. (2003) Position-effect variegation and the genetic dissection of chromatin regulation in *Drosophila*. Semin. Cell. Dev. Biol. **14:** 67–75
17 Jakowitsch J., Papp I., Moscone E. A., Winden J. van der, Matzke M. and Matzke A. J. (1999) Molecular and cytogenetic characterization of a transgene locus that induces silencing and methylation of homologous promoters in trans. Plant J. **17:** 131–140
18 Iglesias V. A., Moscone E. A., Papp I., Neuhuber F., Michalowski S., Phelan T. et al. (1997) Molecular and cytogenetic analyses of stably and unstably expressed transgene loci in tobacco. Plant Cell **9:** 1251–1264
19 Park Y. D., Papp I., Moscone E. A., Iglesias V. A., Vaucheret H., Matzke A. J. et al. (1996) Gene silencing mediated by promoter homology occurs at the level of transcription and results in meiotically heritable alterations in methylation and gene activity. Plant J. **9:** 183–194
20 Hsieh J. and Fire A. (2000) Recognition and silencing of repeated DNA. Annu. Rev. Genet. **34:** 187–204
21 Prols F. and Meyer P. (1992) The methylation patterns of chromosomal integration regions influence gene activity of transferred DNA in *Petunia hybrida*. Plant J. **2:** 465–475

22 Lohuis M. ten, Muller A., Heidmann I., Niedenhof I. and Meyer P. (1995) A repetitive DNA fragment carrying a hot spot for de novo DNA methylation enhances expression variegation in tobacco and petunia. Plant J. **8:** 919–932
23 Koprek T., Rangel S., McElroy D., Louwerse J. D., Williams-Carrier R. E. and Lemaux P. G. (2001) Transposon-mediated single-copy gene delivery leads to increased transgene expression stability in barley. Plant Physiol. **125:** 1354–1362
24 Stam M., Belele C., Dorweiler J. E. and Chandler V. L. (2002) Differential chromatin structure within a tandem array 100 kb upstream of the maize b1 locus is associated with paramutation. Genes Dev. **16:** 1906–1918
25 Goubely C., Arnaud P., Tatout C., Heslop-Harrison J. S. and Deragon J. M. (1999) S1 SINE retroposons are methylated at symmetrical and non-symmetrical positions in *Brassica napus*: identification of a preferred target site for asymmetrical methylation. Plant Mol. Biol. **39:** 243–255
26 Yoder J. A., Walsh C. P. and Bestor T. H. (1997) Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. **13:** 335–340
27 Cockerill P. N. and Garrard W. T. (1986) Chromosomal loop anchorage sites appear to be evolutionarily conserved. FEBS Lett. **204:** 5–7
28 Allen G. C., Spiker S. and Thompson W. F. (2000) Use of matrix attachment regions (MARs) to minimize transgene silencing. Plant Mol. Biol. **43:** 361–376
29 Mankin S. L., Allen G. C., Phelan T., Spiker S. and Thompson W. F. (2003) Elevation of transgene expression level by flanking matrix attachment regions (MAR) is promoter dependent: a study of the interactions of six promoters with the RB7 3' MAR. Transgen. Res. **12:** 3–12
30 Mlynarova L., Hricova A., Loonen A. and Nap J.P. (2003) The presence of a chromatin boundary appears to shield a transgene in tobacco from RNA silencing. Plant Cell **15:** 2203–2217
31 Matzke A. J. and Matzke M. A. (1998) Position effects and epigenetic silencing of plant transgenes. Curr. Opin. Plant Biol. **1:** 142–148
32 Kumpatla S. P., Chandrasekharan M. B., Iyer L. M., Li G. F. and Hall T. C. (1998) Genome intruder scanning and modulation systems and transgene silencing. Trends Plant Sci. **3:** 97–104
33 Wu K., Malik K., Tian L., Hu M., Martin T., Foster E. et al. (2001) Enhancers and core promoter elements are essential for the activity of a cryptic gene activation sequence from tobacco, *tCUP*. Mol. Genet. Genom. **265:** 763–770
34 Okresz L., Mathe C., Horvath E., Schell J., Koncz C. and Szabados L. (1998) T-DNA trapping of a cryptic promoter identifies an ortholog of highly conserved SNZ growth arrest response genes in *Arabidopsis*. Plant Sci. **138:** 217–228
35 Ingelbrecht I., Breyne P., Vancompernolle K., Jacobs A., Van Montagu M. and Depicker A. (1991) Transcriptional interference in transgenic plants. Gene **109:** 239–242
36 Meza T. J., Kamfjord D., Håkelien A. M., Evans I., Godager L. H., Mandal A. et al. (2001) The frequency of silencing in *Arabidopsis thaliana* varies highly between progeny of siblings and can be influenced by environmental factors. Transgenic Res. **10:** 53–67
37 Meza T. J., Enerly E., Børud B., Larsen F., Mandal A., Aalen R. B. et al. (2002) A human CpG island randomly inserted into a plant genome is protected from methylation. Transgenic Res. **11:** 133–142
38 Butenko M. A., Patterson S. E., Grini P. E., Stenvik G. E., Amundsen S. S., Mandal A. et al. (2003) *INFLORESCENCE DEFICIENT IN ABSCISSION* controls floral organ abscission in *Arabidopsis* and identifies a novel family of putative ligands in plants. Plant Cell **15:** 2296–2307
39 Stangeland B., Nestestog R., Grini P. E., Skrbo N., Berg A., Salehian Z. et al. (2005) Molecular analysis of *Arabidopsis* endosperm and embryo promoter trap lines: reporter-gene expression can result from T-DNA insertions in antisense orientation, in introns and in intergenic regions in addition to sense insertion at the 5' end of genes. J. Exp. Bot. **56:** 2495–2505
40 Stangeland B., Salehian Z., Aalen R., Mandal A. and Olsen O.A. (2003) Isolation of GUS marker lines for genes expressed in *Arabidopsis* endosperm, embryo and maternal tissues. J. Exp. Bot. **54:** 279–290
41 Paulin R., Grigg G.W., Davey M.W. and Piper A.A. (1998) Urea improves efficiency of bisulphite-mediated sequencing of 5'-methylcytosine in genomic DNA. Nucleic Acids Res. **26:** 5009–5010
42 Pelissier T., Thalmeir S., Kempe D., Sanger H. L. and Wassenegger M. (1999) Heavy *de novo* methylation at symmetrical and non-symmetrical sites is a hallmark of RNA-directed DNA methylation. Nucleic Acids Res. **27:** 1625–1634
43 Pelissier T. and Wassenegger M. (2000) A DNA target of 30 bp is sufficient for RNA-directed DNA methylation. RNA 6: 55–65
44 Aufsatz W., Mette M. F., Winden J. van der, Matzke A. J. and Matzke M. (2002) RNA-directed DNA methylation in Arabidopsis. Proc. Natl. Acad. Sci. USA **99 (Suppl. 4):** 16499–16506
45 Koncz C. and Schell J. (1986) The promoter of T_L-DNA gene *5* controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. Mol. Gen. Genet. **204:** 383–396
46 Clough S. J. and Bent A. F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:** 735–743
47 Grini P. E., Jürgens G. and Hülskamp M. (2002) Embryo and endosperm development is disrupted in the female gametophytic *capulet* mutants of *Arabidopsis*. Genetics **162:** 1911–1925
48 Yamada K., Lim J., Dale J. M., Chen H., Shinn P., Palm C. J. et al. (2003) Empirical analysis of transcriptional activity in the *Arabidopsis* genome. Science **302:** 842–846
49 Meyers B. C., Tej S. S., Vu T. H., Haudenschild C. D., Agrawal V., Edberg S. B. et al. (2004) The use of MPSS for whole-genome transcriptional analysis in Arabidopsis. Genome Res. **14:** 1641–1653
50 The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796–815
51 Benham C., Kohwi-Shigematsu T. and Bode J. (1997) Stress-induced duplex DNA destabilization in scaffold/matrix attachment regions. J. Mol. Biol. 274: 181–196
52 Frisch M., Frech K., Klingenhoff A., Cartharius K., Liebich I. and Werner T. (2002) *In silico* prediction of scaffold/matrix attachment regions in large genomic sequences. Genome Res. **12:** 349–354
53 Glazko G. V., Rogozin I. B. and Glazkov M. V. (2001) Comparative study and prediction of DNA fragments associated with various elements of the nuclear matrix. Biochim. Biophys. Acta. **1517:** 351–364
54 Singh G. B., Kramer J. A. and Krawetz S. A. (1997) Mathematical model to predict regions of chromatin attachment to the nuclear matrix. Nucleic Acids Res. **25:** 1419–1425
55 Drunen C. M. van, Sewalt R. G., Oosterling R. W., Weisbeek P. J., Smeekens S. C. and Driel R. van (1999) A bipartite sequence element associated with matrix/scaffold attachment regions. Nucleic Acids Res. **27:** 2924–2930
56 Aufsatz W., Mette M. F., Winden J. van der, Matzke M. and Matzke A. J. (2002) HDA6, a putative histone deacetylase needed to enhance DNA methylation induced by double-stranded RNA. EMBO J. **21:** 6832–6841
57 Cao X., Aufsatz W., Zilberman D., Mette M. F., Huang M. S., Matzke M. et al. (2003) Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. Curr. Biol. **13:** 2212–2217
58 Wassenegger M. (2000) RNA-directed DNA methylation. Plant Mol. Biol. **43:** 203–220

59 Morel J. B., Mourrain P., Beclin C. and Vaucheret H. (2000) DNA methylation and chromatin structure affect transcriptional and post-transcriptional transgene silencing in *Arabidopsis*. Curr. Biol. **10:** 1591–1594
60 Kovarik A., Van Houdt H., Holy A. and Depicker A. (2000) Drug-induced hypomethylation of a posttranscriptionally silenced transgene locus of tobacco leads to partial release of silencing. FEBS Lett. **467:** 47–51
61 Kilby N. J., Leyser H. M. and Furner I. J. (1992) Promoter methylation and progressive transgene inactivation in *Arabidopsis*. Plant Mol. Biol. **20:** 103–112
62 Jones L., Ratcliff F. and Baulcombe D. C. (2001) RNA-directed transcriptional gene silencing in plants can be inherited independently of the RNA trigger and requires Met1 for maintenance. Curr. Biol. **11:** 747–757
63 Cao X., Springer N. M., Muszynski M. G., Phillips R. L., Kaeppler S. and Jacobsen S. E. (2000) Conserved plant genes with similarity to mammalian de novo DNA methyltransferases. Proc. Natl.Acad. Sci. USA **97:** 4979–4984
64 Matzke M., Aufsatz W., Kanno T., Daxinger L., Papp I., Mette M. F. et al. (2004) Genetic analysis of RNA-mediated transcriptional gene silencing. Biochim. Biophys. Acta **1677:** 129–141
65 Mette M. F., Winden J. van der, Matzke M. A. and Matzke A. J. (1999) Production of aberrant promoter transcripts contributes to methylation and silencing of unlinked homologous promoters in trans. EMBO J. **18:** 241–248
66 Mette M. F., Aufsatz W., Winden J. van der, Matzke M. A. and Matzke A. J. (2000) Transcriptional silencing and promoter methylation triggered by double-stranded RNA. EMBO J. **19:** 5194–5201
67 Breyne P., Gheysen G., Jacobs A., Van Montagu M. and Depicker A. (1992) Effect of T-DNA configuration on transgene expression. Mol. Gen. Genet. **235:** 389–396
68 Francis K.E. and Spiker S. (2005) Identification of *Arabidopsis thaliana* transformants without selection reveals a high occurrence of silenced T-DNA integrations. Plant J. **41:** 464–477
69 Smale S.T. and Kadonaga J.T. (2003) The RNA polymerase II core promoter. Annu. Rev. Biochem. **72:** 449–479