Planta (2001) 213: 731–740
DOI 10.1007/s004250100535

ORIGINAL ARTICLE

Sandeep Kumar · Matthias Fladung

Gene stability in transgenic aspen (*Populus*). II. Molecular characterization of variable expression of transgene in wild and hybrid aspen

Received: 10 October 2000 / Accepted: 20 December 2000 / Published online: 23 March 2001

Abstract In many annual plant species, transgene inactivation occurs most often when multiple incomplete/complete copies of the transgene are present in a genome. The expression of single-copy transgene loci may also be negatively influenced by the flanking plant DNA and/or chromosomal location (position effect). To understand transgene silencing in a long-lived tree system, we analyzed several wild (*Populus tremula* L.) and hybrid (*P. tremula* L. × *P. tremuloides* Michx.) aspen lines transgenic to the *rolC* phenotypical marker system and grown under in vitro, greenhouse and field conditions. The morphological features of the 35S-*rolC* gene construct were used to screen lines with altered transgene expression, which was later confirmed by Northern experiments. Molecular analyses of hybrid aspen revealed that transgene inactivation was always a consequence of transgene repeats. In wild non-hybrid aspen, however, multiple-insertion-based altered or loss of *rolC* expression was observed only in three out of six lines showing transgene inactivation. Sequencing analysis revealed AT-rich patches at the transgene flanking genomic regions of some of the wild aspen transgenic lines. One wild aspen line showing variable *rolC* expression revealed characteristic integration of the transgene into genomic regions containing a high AT content (85% or more). In the remaining two wild aspen transgenic lines unstable for *rolC* expression, single-copy integration and non-AT-rich or repeat-free transgene flanking regions were found. A partial suppression of *rolC* was observed in some plants of one of the field-grown wild aspen transgenic lines. In the other wild aspen transgenic line an additional mutant phenotype along with transgene inactivation was found. This indicates that the host genome has some control over expression of a transgene, and the possible role of AT-rich regions in defense against foreign DNA.

Keywords *Populus* (transgene) · Position effect · Promoter methylation · T-DNA repeat · Transgene integration · Transgene silencing

Abbreviations Amp^R: gene for ampicillin resistance · DIG: digoxigenin · HDGS: homology-dependent gene silencing · *nptII*: gene for neomycin phosphotransferase II · rpPCR: reverse primer polymerase chain reaction

Introduction

The insertion of foreign DNA into a plant genome may lead to alterations in its structure, which may have effect on host and/or transgene expression. Epigenetic transgene inactivation may occur when multiple copies of a transgene are present in a genome, which is frequently associated with promoter methylation (Hobbs et al. 1990, 1993; Kilby et al. 1992; Assad et al. 1993; Matzke et al. 1994; Mette et al. 1999). Transgene silencing of multiple transgene copies seems to result from the interactions between homologous sequences, and is therefore termed homology-dependent gene silencing (HDGS). The expression of single-copy transgene loci may also be negatively influenced by the flanking plant DNA and/or chromosomal location (Kooter et al. 1999). The actual molecular processes responsible for the epigenetic gene inactivation in plants have not been elucidated yet, in part because several potential silencing mechanisms seem to co-exist in most of the cases analyzed (Cerutti et al. 1997).

Stable expression of transgene(s) is important for commercial use of genetic transformation in long-lived tree species as well as for ecological risk-assessment studies. However, analysis of the instable/stable transgene expression in trees is more problematic than in crop plants (Fladung et al. 1997). To our knowledge, little is known about the silencing of transgenes in forest trees

S. Kumar (✉) · M. Fladung
Bundesforschungsanstalt für Forst- und Holzwirtschaft,
Institut für Forstgenetik und Forstpflanzenzüchtung,
Sieker Landstrasse 2, 22927 Großhansdorf, Germany
E-mail: kumar@holz.uni-hamburg.de
Fax: +49-4102-696200

(Fladung 1999). Trees having extended vegetative phases and genomes which are undisturbed by human manipulations offer an unique opportunity to study and understand the molecular mechanisms of transgene inactivation which may remain unnoticed in short-term annual crops.

In this report, transgenic wild and hybrid aspen carrying a morphologically visible 35S-*rolC* gene construct (Fladung et al. 1997) were analyzed to establish a molecular basis for instable/stable transgene expression. The structure of integrated transgenes and the flanking genomic sequences were analyzed in different aspen transgenic lines characterized by stable and unstable *rolC* expression.

Materials and methods

Plant materials, transformation, cultivation and morphological analysis

A leaf-disc co-cultivation methodology was used for the *Agrobacterium*-mediated binary vector transformation of two wild non-hybrid aspen [*Populus tremula* L.: 'Brauna11', Kamez (Saxon), Germany; 'W52', Tapian, Russia] clones, and one hybrid aspen 'Esch5' (*P. tremula* L. clone 'Brauna11' × *P. tremuloides* Michx. Clone 'Tur 141': Woodstock, N. H., USA) clone (Fladung et al. 1997). The plasmid pPCV002-CaMVC carries the chimeric *rolC* gene from the Ri-plasmid of *Agrobacterium rhizogenes* under the control of the 35S-promoter of the cauliflower mosaic virus (Spena et al. 1987).

Transgenic plants propagated in vitro were transferred to, and cultivated in, a greenhouse under natural daylight conditions (Fladung et al. 1997; Fladung 1999). Aspen plants transgenic for the 35S-*rolC* construct are characterized by a reduction in plant height associated with shortened internodes, smaller leaves, and lower chlorophyll contents in leaves, as also described in transgenic tobacco and potato (Spena et al. 1987; Fladung 1990; Fladung et al. 1996, 1997). For the present study, 15 hybrid aspen-based transgenic lines and 16 lines regenerated from the wild aspen (Fladung et al. 1997) were cultured in vitro and monitored for visible *rolC* expression. All the Esch5 hybrid aspen-based transgenic lines were transferred to the greenhouse. Four Esch5-based lines namely, Esch5:35S-*rolC*#1, #3, #5, and #16, Southern blotted to contain a single *rolC* copy (Fladung et al. 1997) were transferred to the field. Similarly, five wild aspen transgenic lines (out of the seven lines that could survive long-term in vitro culturing) were transferred to the greenhouse. The lines transferred were: W52:35S-*rolC*#2, #9, and #3; and Brauna11:35S-*rolC*#5, and #2. Only two wild aspen transgenic lines (W52:35S-*rolC*#2 and Brauna11:35S-*rolC*#2) were field-planted (Fladung and Muhs 1999). The transgenic plants were compared to non-transformed control plants for the morphological expression of the *rolC* gene during their continuous growth for 5–6 years in the greenhouse and the field.

Analyses of DNA and RNA, determination of T-DNA repeats, and promoter methylation

Extraction of genomic DNA and total RNA from leaves for polymerase chain reaction (PCR), Southern and Northern analyses, enzymatic restriction of DNA, electrophoresis, and blotting of DNA and RNA onto Biodyne A membranes (Pall Europe, Portsmouth, UK) were done as described by Fladung et al. (1996, 1997). The non-radioactive digoxigenin (DIG) system used for pre-hybridization and hybridization utilizing DIG-dUTP-labeled purified *rolC* in Southern and Northern experiments has been described previously (Fladung and Ahuja 1995; Fladung et al. 1997; Fladung 1999).

Determination of T-DNA repeats was done using the reverse primer PCR (rpPCR) method (Kumar and Fladung 2000a) which utilizes different primer pairs oriented in opposite directions in the construct (Fig. 4A) so that no amplification is possible in the case of single-copy integration. However, amplification products will be obtained when the transgene is integrated in the form of multiple repeats (Fig. 4B). Additional primer pairs were used to detect formation of inverted transgenes repeats (Fig. 4C).

Transgene promoter methylation was studied by restriction of approximately 1 μg of genomic DNA with the methylation-sensitive enzymes *Hpa*II or *Msp*I, which are differentially sensitive to internal cytosine methylation at the sequence CCGG. The suitability of the isoschizomeric pair *Hpa*II/*Msp*I for PCR-based detection of promoter methylation at a key position has recently been reported (Fu et al 2000; Kumar and Fladung 2000a). The restriction reactions were performed in a 10-μl volume, using manufacturer's buffer (Roche Molecular Biochemicals, Mannheim, Germany), incubated overnight at 37 °C to ensure complete digestion when the restriction site is not methylated. The restricted DNA was subjected to PCR amplification using primers specifically designed to amplify the promoter region, positive amplification indicating methylation of the promoter at the restriction site. For a positive control, a primer pair was added for the amplification of a small genomic region that had no restriction site for the enzymes used (Kumar and Fladung 2000a).

Sequencing of T-DNA repeat junctions and genomic regions flanking the transgene

Inverse-PCR was performed as described previously (Fladung 1999) to amplify DNA regions flanking the right and left borders of the T-DNA of transformed aspen lines. The amplified fragments from the T-DNA repeat junctions and inverse-PCR were purified using the High Pure PCR Product Purification Kit (Roche Molecular Biochemicals) according to the manufacturer's recommendations. Sequencing of the purified fragments was done in a PCR reaction using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, Foster City, Calif., USA) as described previously (Fladung 1999). Unless stated otherwise, the flanking genomic sequences of 500 base pairs (bp) adjacent to the left and right T-DNA borders were analyzed. The percent A + T base content (% AT) was calculated for the genomic sequences obtained. The genomic regions having more than a 70% AT value have been described as AT-rich (Gallie 1998). In this report, the regions of minimum length 50 bp having an AT value of 85% or more were marked and the sequence arrangement of average AT value (60–65%) with some patches (50 bp or more) of 70–80% was termed "normal" nucleotide distribution.

Results

Regeneration, cultivation and morphological analyses of transgenic wild/hybrid aspen

We have previously reported (Fladung et al. 1997) that by using *Agrobacterium tumefaciens* strain LBA4404, only the hybrid aspen clone Esch5 could be successfully regenerated on kanamycin-containing medium and no regeneration could be obtained in any of the wild aspen clones. Although, transgenic wild aspen plants were obtained using the GV3101 strain, the regeneration frequency was low except for the clone W52 for which an improved transformation method was used (Fladung et al. 1997). Further, despite the higher initial

transformation success in the W52 wild aspen clone, a high subsequent in vitro mortality was observed.

Under in vitro conditions, long-term (5–6 years) morphological observations of *rolC* expression have so far revealed a stable *rolC* phenotype in all the 15 hybrid aspen transgenic lines. Among 16 wild aspen transgenic lines, only seven lines could survive during the long-term in vitro culturing. Morphological reversions with incomplete/complete suppression of *rolC* phenotypic expression were observed in some of the in vitro-maintained wild aspen transgenic lines. In the first year of in vitro cultivation, the phenotypic features of the transgenic line W52:35S-*rolC*#9 were comparable to those of the control plants (Fladung et al. 1997). Other wild aspen transgenic lines (W52:35S-*rolC*#2, W52:35S-*rolC*#12) showed variation in *rolC* expression after 4–5 years of in vitro cultivation.

Major alterations in *rolC* expression in most of the transgenic lines, however, were observed after transfer from in vitro conditions to greenhouse or field. Incomplete/complete suppression of transgene expression was observed in three Esch5-based transgenic lines (Esch5:35S-*rolC*#1, Esch5:35S-*rolC*#2 and Esch5:35S-*rolC*#12) out of the 15 lines transferred to the greenhouse. Among five wild aspen transgenic lines transferred to the greenhouse, three lines (Brauna11:35S-*rolC*#5, W52:35S-*rolC*#9 and #3) were observed with altered or reverted transgene expression. Two of the field-planted transgenic lines, one each from hybrid and wild aspen (Esch5:35S-*rolC*#1 and Brauna11:35S-*rolC*#2, respectively) showed variable transgene expression. We did not observe any transgene inactivation in the other three field-grown hybrid aspen lines (Esch5:35S-*rolC*#3, #5, and #16). Similarly, unaltered morphological *rolC* features were found in the one remaining field-planted wild aspen transgenic line (W52:35S-*rolC*#2) – the line which showed variable in vitro transgene expression. However, only a limited number of aspen transgenic lines could be transferred to the field and the possibility of other lines turning out to be unstable after field transfer is not ruled out.

Reversion of the *rolC* phenotype to wild type in the line Esch5:35S-*rolC*#1 has already been reported (Fig. 1A; Fladung 1999). The phenotypically visible *rolC* expression in two other hybrid aspen-based transgenic lines (Esch5:35S-*rolC*#2 and Esch5:35S-*rolC*#12) decreased gradually over a period of 3–4 years of cultivation in the greenhouse. Subsequently, the transgene expression was completely suppressed after 4–5 years and the phenotypes of these lines are presently comparable to the control plants. The loss of *rolC* expression seemed to be stable in these lines and the plants, once reverted, maintained the changed features in the following years (Table 1). A similar stable complete *rolC* suppression was also observed in the two W52-based wild aspen transgenic lines (W52:35S-*rolC*#9 and W52:35S-*rolC*#3). Leaves of the line W52:35S-*rolC*#3, however, were found to be morphologically different both from the control and *rolC* phenotype (Fig. 1B).

Compared to the lines showing completely suppressed *rolC* phenotypes, the alterations in morphological expression of the transgene were more complex and variable in the other wild aspen-based transgenic lines (W52:35S-*rolC*#2, Brauna11:35S-*rolC*#2 and Brauna11:35S-*rolC*#5). The transgenic line Brauna11:35S-*rolC*#5 frequently showed reverted shoots with large internodes and large leaves (Fig. 2A) and, interestingly, re-reversion back to the *rolC* phenotype (data not shown). Incomplete *rolC* suppression in some plants of the transgenic line Brauna11:35S-*rolC*#2 has, so far, only been observed under field conditions. The altered plants from this transgenic line showed morphological features different from both the control and *rolC* phenotype

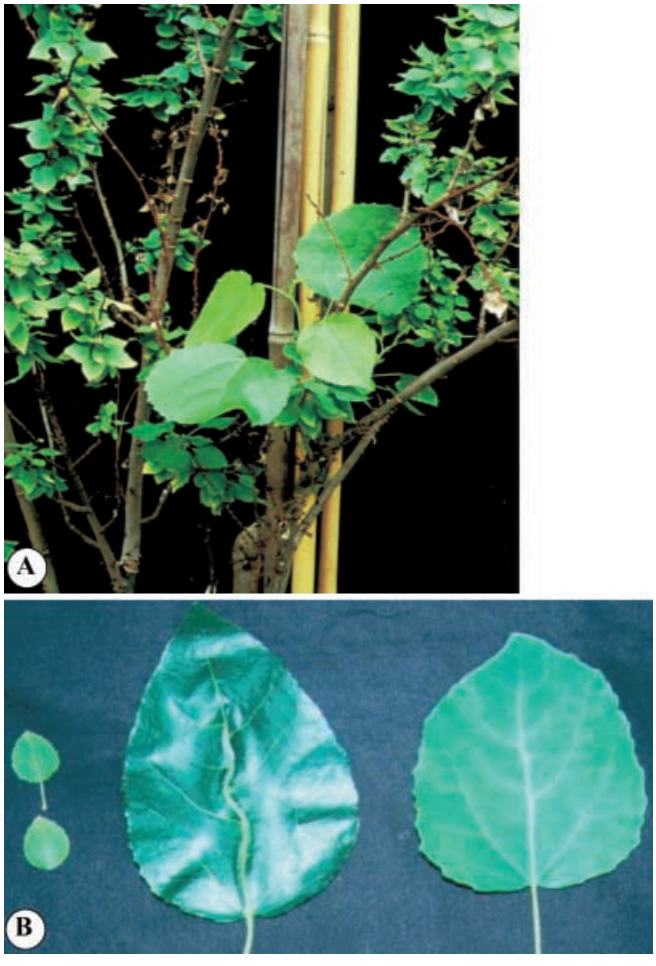


Fig. 1A, B Morphological reversion in wild (*Populus tremula*) and hybrid (*P. tremula* × *P. tremuloides*) aspen transformed with the *rolC* gene. **A** An example of variable *rolC* expression (see also Table 1) in an unstable aspen transgenic line showing phenotypic reversion in a side branch (with large leaves). The rest of the plant shows the *rolC* phenotype. This figure represents the situation observed in the transgenic lines Brauna11:35S-*rolC*#5 and Esch5:35S-*rolC*#1 (Fladung 1999). **B** Comparison of a wild-type leaf (*right*), a leaf from transgenic line W52:35S-*rolC*#3 (*middle*), and leaves from Esch5 hybrid aspen-based transgenic line (Esch5:35S-*rolC*#3) showing a stable *rolC* phenotype (*left*). Note the glossy phenotype of the leaf from the transgenic line W52:35S-*rolC*#3

Table 1 *RolC* expression and transgene structure in different transgenic lines of wild (*Populus tremula*) and hybrid (*P. tremula* × *P. tremuloides*) aspen

Wild/hybrid aspen	Transgenic line	Southern analysis[a]	T-DNA structure (rpPCR)[b]	Promoter methylation[c]	*RolC* phenotypic expression[d]
Hybrid aspen	Esch5:35S-*rolC*#3, #4, #5, #16	One copy	No repeat	NM	*rolC* type
	Esch5:35S-*rolC*#1	One copy	Inverted repeat (one incomplete)	M	Variable
	Esch5:35S-*rolC*#2	Two copies	Direct repeat (both complete)	M	Reverted
	Esch5:35S-*rolC*#12	Two copies	Direct repeat (both complete)	M	Reverted
Wild aspen	W52:35S-*rolC*#9	One copy	Direct repeat (one incomplete)	M	Reverted
	W52:35S-*rolC*#12[e]	Two copies	Direct repeat (both incomplete)	M	Variable
	W52:35S-*rolC*#2	One copy	No repeat	NM	Variable[f]
	W52:35S-*rolC*#3	One copy	No repeat	NM	Reverted[g]
	Brauna11:35S-*rolC*#2	One copy	No repeat	NM	Altered
	Brauna11:35S-*rolC*#5	Two copies	Inverted repeat (one incomplete)	M	Variable

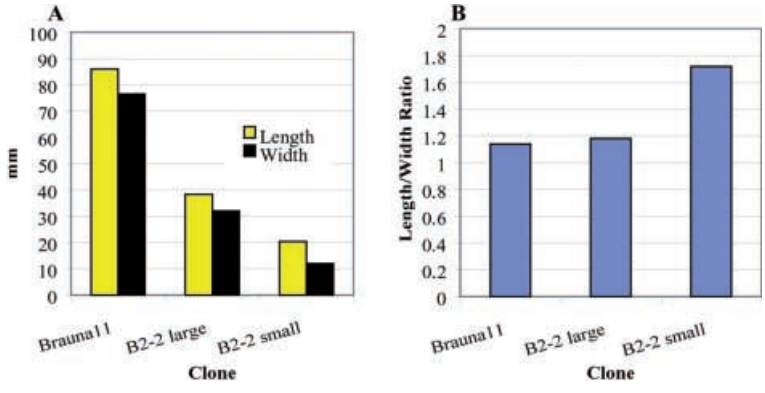
[a]From Fladung et al. 1997 (*rolC* probed)
[b]Kumar and Fladung 2000a
[c]Cytosine methylation at the sequence CCGG (Kumar and Fladung 2000a); *M* methylated, *NM* not methylated
[d]*Variable* indicates *rolC* expression is suppressed only in one or two branches/leaves of a plant; *Reverted* indicates *rolC* expression is completely suppressed in the whole plant; *Altered* indicates incomplete *rolC* suppression in the whole plant
[e]No plant transferred to greenhouse or field
[f]Alteration in *rolC* expression observed in in vitro cultures only
[g]Additionally mutant phenotype

(Fig. 2A, B). The length and width of the leaves collected from the reverted plants were in between those of the control and 35S-*rolC* phenotype (Fig. 2A). However, the length/width ratio of the reverted leaves was comparable to that of the control plants (Fig. 2B). The reverted morphological expression was confirmed by Northern experiments, which clearly showed very weak *rolC*-specific transcripts from the leaves of reverted plants grown under field conditions. The *rolC*-specific transcript was, however, present in leaves collected from the plants maintained in the greenhouse and from *rolC*-expressing plants in the field (Fig. 3). The change in *rolC* expression was most variable in the transgenic line W52:35S-*rolC*#2; moreover, the suppression was observed only in the leaves of in vitro cultures and was reversible in the subsequent subculturing. We have, so far, not observed any appreciable morphological change in this transgenic line under greenhouse or field conditions.

Number and structure of integrated transgenes, and their promoter methylation

The data pertaining to Southern analysis of different transgenic lines analyzed in the present study are shown in Fig. 4E and summarized in Table 1. A number of transgenic lines were identified showing the insertion of one copy of the *rolC* gene. Interestingly, all the

Fig. 2A, B Length and width measurements of aspen leaves collected from the altered (*B2-2 large*, see also Table 1) and *rolC*-expressing (*B2-2 small*) plants of transgenic line Brauna11:35S-*rolC*#2 compared to leaves from a non-transformed control (*Brauna11*). **A** The leaves collected from the altered plants are smaller in length and width compared to the leaves from the control plants. However, the reverted leaves are larger when compared to the leaves collected from the *rolC*-expressing transgenic plants. **B** The length/width ratio of leaves from the reverted transgenic plants is comparable to the ratio of leaves collected from the control plants



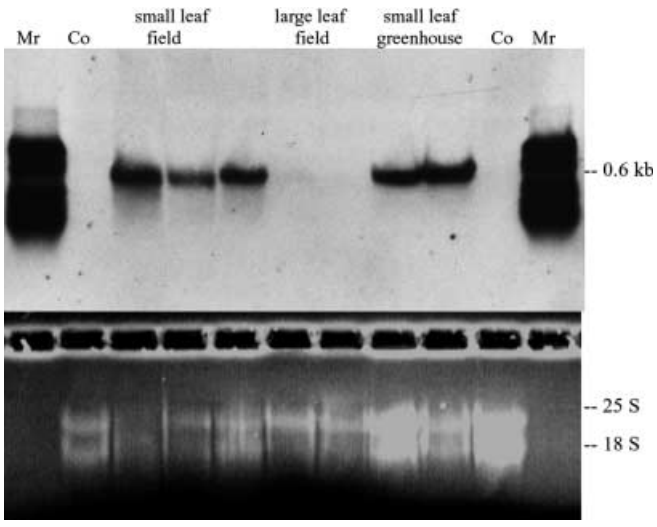


Fig. 3 Northern analysis of RNA (*upper part*) extracted from 35S-*rolC*-expressing small leaves growing in greenhouse and field, and large leaves from field-grown plants with altered *rolC* expression in the transgenic aspen line Brauna11:35S-*rolC*#2. Total RNA was probed with DIG-labeled *rolC*. No signal was observed in the non-transformed aspen (*Co*), and a very weak signal was obtained in the large leaves collected from the field-grown plants showing altered *rolC* expression. *Mr* DIG-labeled DNA molecular weight marker II (Roche Molecular Biochemicals, Mannheim, Germany)

transgenic lines observed to contain two *rolC* copies in Southern experiments were found positive to transgene repeat analysis using rpPCR (Table 1; Fig. 4E and F, lanes 8 and 9, respectively). It implies that in transgenic aspen lines where more than one T-DNA copy is integrated, in most cases it is in linked form (Kumar and Fladung 2000b). Surprisingly, two transgenic lines (Esch5:35S-*rolC*#1 and W52:35S-*rolC*#9) observed by Southern blotting to harbor a single *rolC* copy were also found to contain incomplete T-DNA repeats (Fig. 4E and F, lanes 2 and 7, respectively). The sequencing of the T-DNA repeat junctions confirmed the rpPCR results for the repeat formation. Figure 5 shows the structures of the integrated T-DNA(s) in the transgenic lines containing direct repeats. In the transgenic lines Esch5:35S-*rolC*#1 and Brauna11:35S-*rolC*#5, amplification fragments were obtained (Fig. 4F, lanes 1 and 8, respectively) using primer pair 2 + 3 (Fig. 4C), which were shorter than the size expected for a complete inverted repeat (Fig. 4C). No amplification was obtained with other primer pairs used to detect direct or inverted repeats in these two transgenic lines. These observations indicated a truncated inverted repeat formation and the sequencing of the repeat junctions confirmed this notion. The sequencing results revealed that the left border of the complete T-DNA copy was linked to the neomycin phosphotransferase II gene (*nptII*) and the nopaline synthase promoter (pnos) of the second copy (data not shown) in the transgenic lines Esch5:35S-*rolC*#1 and Brauna11:35S-*rolC*#5, respectively, with the deletion of 35S-*rolC* and the pg5-Amp^R (ampicillin resistance) gene. The structure of the integrated transgenes in the line Esch5:35S-*rolC*#1 was also described previously (Fladung 1999). The transgenic line Brauna11:35S-*rolC*#5 showed an incomplete inverted repeat lacking *rolC* in rpPCR analysis (Fig. 4F, lane 8) and two bands in *rolC*-probed Southern analysis (Fig. 4E, lane 8). Therefore, two independent transgene integration sites exist in this transgenic line and one site contains an incomplete inverted repeat.

The unstable expression of the *rolC* gene on the basis of morphological screenings has already been reported in the transgenic aspen line Esch5:35S-*rolC*#1, which was observed to contain an inverted incomplete T-DNA repeat adjacent to the left border (Fladung 1999). In four so far stable *rolC*-expressing lines, simple structured single-copy transgene insertion loci without methylation of the *Hpa*II/*Msp*I promoter site were found (Fig. 5B, Table 1). As a further demonstration of the T-DNA repeat-related gene silencing, a direct incomplete/complete T-DNA repeat was observed in each of the two different independent hybrid aspen-based transgenic lines (Esch5:35S-*rolC*#2 and #12) (Table 1, Fig. 4F). Similarities between the integrated T-DNAs structures (Fig. 5C, D) indicate that transgene silencing in these two unstable transgenic lines of the hybrid aspen clone is due to the complex repeated T-DNA. Incomplete T-DNA repeats were also observed in three wild aspen-based transgenic lines (W52:35S-*rolC*#9, #12 and Brauna11:35S-*rolC*#5) observed with altered *rolC* morphological expression (Fig. 4F, and Fig. 5E, F). The promoter methylation analysis indicated that at least one of the 35S-promoters was methylated at the *Hpa*II/*Msp*I site in the wild and hybrid aspen transgenic lines containing T-DNA repeats (Table 1). Surprisingly, in three other wild aspen-based transgenic lines showing variable *rolC* phenotype (W52:35S-*rolC*#2, #3 and Brauna11:35S-*rolC*#2) the T-DNA structures were found to be as simple as in the hybrid aspen-based stable lines (Fig. 5B). Thus, despite the single-copy integration, these transgenic wild aspen lines showed altered *rolC* expression under in vitro, greenhouse or field conditions (Table 1).

Sequencing analysis of genomic DNA flanking transgene

The analysis of T-DNA flanking genomic regions in the different wild and hybrid aspen transgenic lines in general revealed an average AT value of 60–65%. However, regions (50 bp or more in length) with a 70–80% AT value were frequently observed in most of the lines analyzed. In this report, the sequence arrangement of an average AT value of 60–65% with some patches of 70–80% is termed a "normal" nucleotide distribution. In three hybrid aspen-based transgenic lines (Esch5:35S-*rolC*#3, #4 and #5), which were marked as stable, this "normal" nucleotide distribution was found over T-DNA flanking genomic regions (Fig. 6). One hybrid aspen transgenic line (Esch5:35S-*rolC*#16) which, so far,

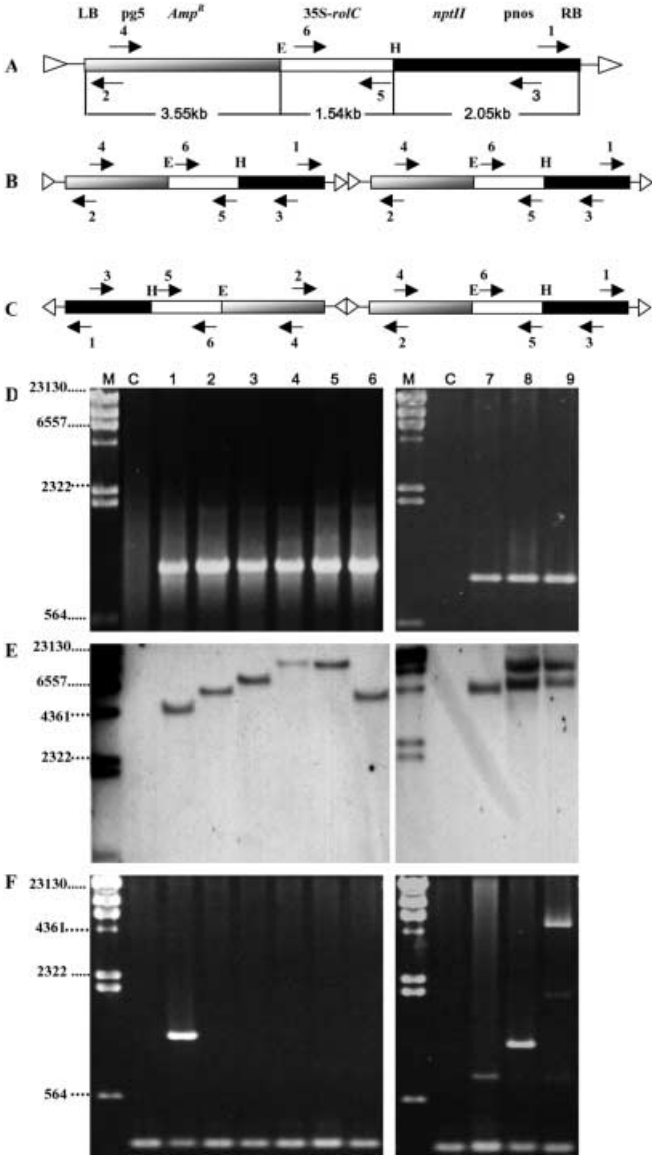


Fig. 4A–F T-DNA of vector pPCV002; and representative PCR, Southern, and rpPCR (Kumar and Fladung 2000a) analyses for different transgenic aspen lines. **A** Schematic diagram showing the T-DNA region with left (*LB*) and right (*RB*) borders of vector pPCV002 containing the 35S-*rolC* gene inserted between *Eco*R1 (*E*) and *Hin*dIII (*H*) sites used for the aspen transformation in this study. The neomycin phosphotransferase II gene (*nptII*), which is controlled by the nos promoter (*pnos*) encoding resistance to kanamycin, is located at the right border and the promoter of gene 5 (*pg5*) and ampicillin resistance gene (Amp^R) are at the left border of the construct. The numbers (*1–6*) represent the locations of different primers used to determine possible transgene repeat formation and the *arrowheads* indicate their respective directions (5′-3′). **B** A direct T-DNA complete repeat formation in head-to-tail integration. Amplification products of expected sizes using five primer pairs (1 + 2, 1 + 5, 1 + 3, 2 + 6, and 2 + 4) will confirm the presence of a complete T-DNA direct repeat. **C** An inverted T-DNA repeat formation in tail-to-tail orientation, which can be detected using three primer pairs (2 + 5, 2 + 3, and 5 + 3). Similarly, an inverted repeat in head-to-head orientation may be determined by another set of primer pairs (1 + 6, 1 + 4, and 6 + 4). **D–F** *Lanes 1–9* Transgenic lines Esch5:35S-*rolC*#1, #3, #5, #16, Brauna11:35S-*rolC*#2, W52:35S-*rolC*#2, W52:35S-*rolC*#9, Brauna11:35S-*rolC*#5, and Esch5:35S-*rolC*#2, respectively. *M* Molecular weight marker II (Roche Molecular Biochemicals, Mannheim). *C* Non-transformed aspen control. **D** Amplification of the *rolC* gene using primer pair 5 + 6. Fragments of correct size were amplified in nine transgenic lines. PCR products were analyzed on a 1.5% agarose gel. **E** Southern blot analysis of genomic DNA extracted from the same nine independent aspen transgenic lines. The genomic DNA was digested with the restriction enzyme *Eco*RI and hybridized to DIG-labeled *rolC* probe. **F** rpPCR analysis of the nine aspen transgenic lines. Fragments amplified using primer pair 1 + 3 for the direct transgene repeats (see **B**) in W52:35S-*rolC*#9 (*lane 7*) and Esch5:35S-*rolC*#2 (*lane 9*) are shown. Amplifications were obtained using primer pair 2 + 3 for the inverted repeats in tail-to-tail orientation (see **C**) in the transgenic lines Esch5:35S-*rolC*#1 (*lane 1*) and Brauna11:35S-*rolC*#5 (*lane 8*). No amplification was obtained in other transgenic lines shown in this figure using different primer pairs to detect direct or inverted repeats. Amplified fragments of less than expected sizes in lanes 1, 7, and 8 suggest that at least one of the T-DNA copies is incomplete in the transgenic lines Esch5:35S-*rolC*#1, W52:35S-*rolC*#9 and Brauna11:35S-*rolC*#5. The band of approximately 100 bp in all the lanes indicates the positive control amplified using aspen genomic primers

is also classified as morphologically stable was, however, found to be flanked by short AT-rich regions (90% and 93%, respectively) on left and right borders (Fig. 6). Among wild aspen transgenic lines, a 260-bp left border flanking genomic sequence (out of 275 bp analyzed) of the line W52:35S-*rolC*#2 was found to have an 88% AT value. Similarly, the transgenic line W52:35S-*rolC*#9 was also flanked by AT-rich regions (about a 180-bp region with a 96% AT value) on the left T-DNA borders (Fig. 6). Three short AT-rich regions (> 50 bp length) with an 86–92% AT value were also observed in one of the right border flanking genomic regions of the transgenic line Brauna11:35S-*rolC*#5. However, we did not observe any AT-rich or DNA-repeat regions in the transgenic lines Brauna11:35S-*rolC*#2 and W52:35S-*rolC*#3, which showed partial *rolC* suppression in some of the field-grown plants and a mutant phenotype in the greenhouse, respectively. Hence, despite the single-copy transgene and non-AT-rich or repeat-free flanking regions, both of these transgenic lines are showing altered or reverted *rolC* morphological expression. Similarly, flanking genomic regions of 80% AT value were also found in three, hybrid aspen-based transgenic lines (out of three analyzed) transformed with rbcS-*rolC* and 35S-*Ac*-*rolC* gene constructs (data not shown). However, *rolC* expression in these lines was restricted only to the leaves and morphological instability/stability was not as conspicuous as in 35S-*rolC* transgenic lines.

Discussion

In this report, we utilized the *rolC* gene as a morphological marker to study the variable transgene expression in wild/hybrid aspen transgenic lines. The transgenic lines were monitored for visible *rolC* expression for a period of 5–6 years during their continuous growth under in vitro, greenhouse and field conditions. The morphological features were later substantiated with

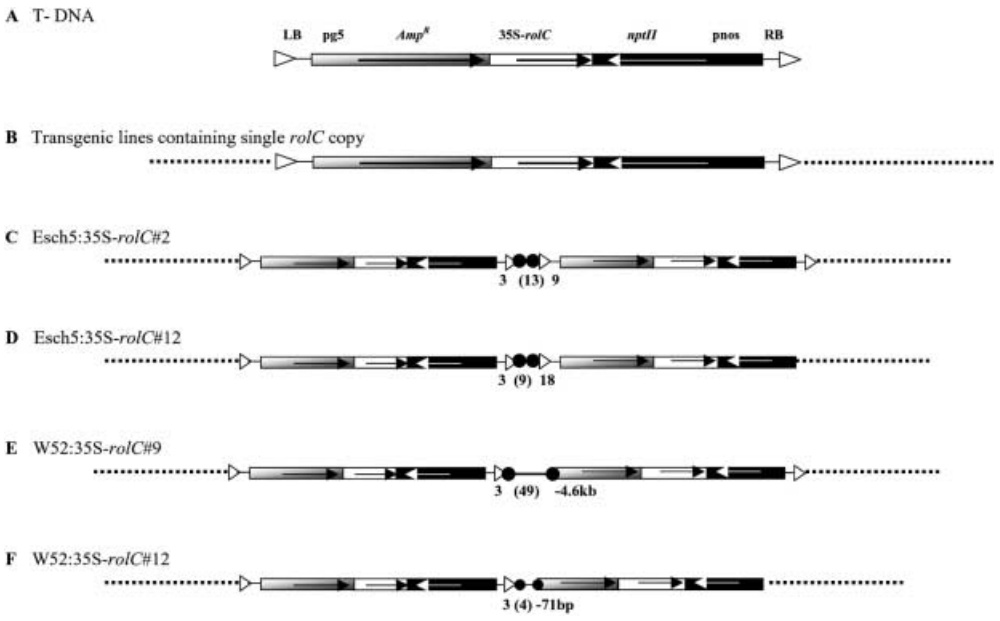


Fig. 5A–F Structure of integrated transgene loci in different aspen and hybrid aspen transgenic lines. The *arrows* indicate 5′-3′orientation of the genes. **A** T-DNA (see Fig. 1A). **B** Single-copy transgene in stable Esch5-based (Esch5:35S-*rolC*#3, #4, #5, and #16) and some of the wild aspen-based (W52:35S-*rolC*#2, #3, and Brauna11:35S-*rolC*#2) transgenic lines. **C–F** Complex T-DNA structure with a complete or incomplete direct repeat in unstable aspen transgenic lines Esch5:35S-*rolC*#2, Esch5:35S-*rolC*#12, W52:35S-*rolC*#9, and W52:35S-*rolC*#12, respectively. The right border (*RB*) of the T-DNA repeat in transgenic lines containing an incomplete repeat is missing (**D, F**). Note the conserved 3-bp RB residual sequence at the repeat junctions of all the four transgenic lines containing a transgene repeat. The RB residual sequence at the repeat junctions is followed by filler DNA (*black dumbbell*) of different lengths, which is interposed between two linked T-DNAs. The left border (*LB*) residual sequence of 9 bp and 18 bp from the T-DNA repeat is present in transgenic lines Esch5:35S-*rolC*#2 and Esch5:35S-*rolC*#12, respectively (**C, D**, respectively). However, in transgenic lines W52:35S-*rolC*#9 and W52:35S-*rolC*#12 (**E, F**, respectively), 4.6 kb and 71 bp, respectively, from the T-DNA repeat are deleted along with LB. The transgenic lines shown in **C**, **E** and **F** have already been described for the structure of integrated transgenes (Kumar and Fladung 2000a)

molecular analyses to determine and establish a molecular basis for transgene silencing in wild/hybrid aspen. The comparison of Southern and morphological analyses of different aspen transgenic lines in the present study indicated that the lines with two-copy transgene insertions turned out to be unstable for *rolC* expression in different stages of their growth (Table 1). Similarly, Hobbs et al. (1990, 1993) also observed reduced transgene expression with increased copy number in transgenic tobacco. In order to determine if the T-DNAs in the transgenic lines containing two copies are inserted in linked form we conducted a rpPCR analysis (Kumar and Fladung 2000a). It revealed that the T-DNAs were inserted in the form of repeats in all the transgenic lines observed with two copies of the transgene (Table 1). Interestingly, two transgenic lines shown to harbor a single transgene (*rolC*) copy in Southern experiments were found to contain multiple T-DNA insertions using rpPCR (Esch5:35S-*rolC*#1 and W52:35S-*rolC*#9). This is understandable because Southern analyses in these lines were conducted on *Hin*dIII- and *Eco*RI-restricted genomic DNA using *rolC* as a probe (Fladung et al. 1997). The *Hin*dIII site was probably methylated in these transgenic lines and the second copy of the transgene was integrated in truncated form and did not contain any *Eco*RI site (Fig. 1A). Similarly, in Southern analysis of transgenic petunia, Cluster et al. (1996) have also reported that a plant genomic *Hin*dIII site was modified in eight transformants.

The transgene-silencing phenomenon has previously been distinguished in two general categories. The first category is termed a position effect, in which the flanking plant DNA and/or chromosomal location negatively influence the expression of the single transgene loci. The second class of transgene inactivation represents the homology-dependent gene silencing (HDGS) and occurs when multiple copies of a specific sequence are present in a genome. Depending on the level at which silencing occurs, HDGS has been distinguished into transcriptional gene silencing (TGS) and post-transcriptional gene silencing (PTGS). Transcriptionally silenced transgenes acquire metastable epigenetic states that are characterized by altered methylation patterns and chromatin structure. Reversible promoter methylation may be used as a useful marker for TGS (Kooter et al. 1999). Considering the stable alteration of the *rolC* expression associated with *Hpa*II/*Msp*I site methylation of at least one of the two 35S promoters in different aspen transgenic lines containing T-DNA repeats in the present study (Table 1), the transgene silencing in these lines may be attributable to HDGS-based TGS. Silencing based on repeated/rearranged transgenes has previously been described in crop plants (Flavell 1994; Kooter et al. 1999) which mainly includes *Arabidopsis* (Kilby et al.

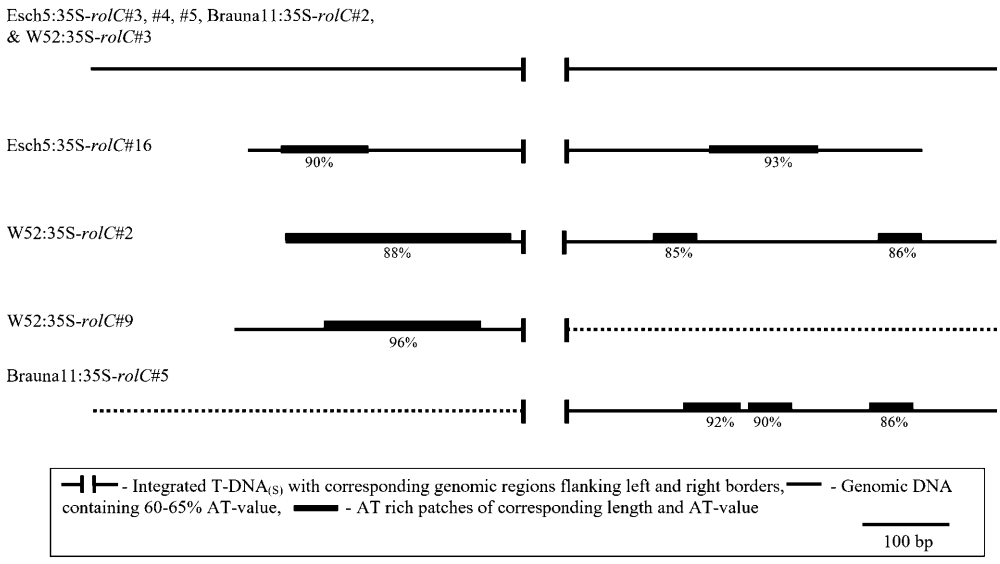


Fig. 6 Sequencing analysis of T-DNA(s) flanking genomic sequences of different aspen transgenic lines. The integrated T-DNA(s) is shown as two parallel bars in the middle and the lines on the left and right of these bars indicate left and right border flanking genomic sequences. Unless stated otherwise, the length of the line corresponds to the number of base pairs analyzed in different transgenic lines (see 100-bp scale). The lengths and the corresponding EMBL accession numbers of flanking genomic sequences analyzed (left and right border, respectively) in the transgenic lines without AT-rich regions (printed at the top of the figure) are: 500 bp (AJ005287) and 197 bp (AJ005290) for Esch5:35S-*rolC*#3; 470 bp (AJ296081) for Esch5:35S-*rolC*#4 (only right border); 492 bp (AJ005756) and 318 bp (AJ005291) for Esch5:35S-*rolC*#5; 51 bp (AJ296083) and 500 bp (AJ296084) for Brauna11:35S-*rolC*#2; 500 bp (AJ296082) and 500 bp (AJ286128) for W52:35S-*rolC*#3. The EMBL accession numbers for the other lines are as follows: Esch5:35S-*rolC*#16 left and right borders, AJ005288 and AJ005292, respectively; W52:35S-*rolC*#2 left and right borders, AJ286126 and AJ286125, respectively; W52:35S-*rolC*#9 left border, AJ286132; Brauna11:35S-*rolC*#5 right border, AJ291296

1992; Assad et al. 1993), tobacco (Hobbs et al. 1990, 1993; Matzke et al. 1994; Park et al. 1996; Krizkova and Hrouda 1998), petunia (Linn et al. 1990; Stam et al. 1997b), and rice (Kohli et al. 1999; Morino et al. 1999).

We did not observe any transgene repeats in the other three wild aspen-based transgenic lines unstable for *rolC* expression (W52:35S-*rolC*#2, W52:35S-*rolC*#3, and Brauna11:35S-*rolC*#2). The T-DNA-flanking genomic regions were sequenced to screen for the unusual nucleotide distribution in the flanking genomic regions in order to explain their altered or variable *rolC* expression. The region flanking the left border in the transgenic line W52:35S-*rolC*#2 was found to be rich in AT content (88% AT value). Similarly AT-rich regions (86–96% AT value) were also observed in left and right borders of transgenic lines W52:35S-*rolC*#9 and Brauna11:35S-*rolC*#5, respectively. This was in sharp contrast to the non-AT-rich flanking regions with "normal" nucleotide distribution in three hybrid aspen-based transgenic lines (Esch5:35S-*rolC*#3, #4 and #5) marked stable for *rolC* expression (Fladung 1999). The high AT values observed in some wild aspen transgenic lines clearly differ from the compositional distribution reported for the transcribed chromosomal regions in dicots which show an AT content around 54% (Salinas et al. 1988). AT-rich and/or tandem repetitive DNA sequences are major DNA components of "constitutive heterochromatin" which is usually found in pericentromeric regions (Hankeln et al. 1996). Euchromatic genes brought into juxtaposition with heterochromatin may exhibit 'position effect variegation' (Cryderman et al. 1999). Although not proven, the variable in vitro *rolC* expression observed in the transgenic line W52:35S-*rolC*#2 containing a single copy of the transgene may be as a result of integration of the transgene near to heterochromatin in this line. The transgenic lines Brauna11:35S-*rolC*#5 and W52:35S-*rolC*#9 were found to contain multiple transgenes and AT-rich flanking regions as well.

The transgenic lines Brauna11:35S-*rolC*#2 and W52:35S-*rol*C#3 showed variable *rolC* expression despite single-copy transgene integration and non-AT-rich flanking genomic regions. Considering the mutant phenotype of the transgenic line W52:35S-*rol*C#3, it may be speculated that the transgene has integrated into a transcribed region resulting in silencing of both the transgene and the endogenous gene. This assumption, however, needs to be proven and detailed investigation of this line is in progress. In the transgenic line Brauna11:35S-*rolC*#2, Northern analysis and the morphological features clearly indicated the alteration in *rolC* transcription and expression. However, the molecular analyses revealed single-copy transgene integration. The AT-base content of transgene flanking genomic sequences in this line is comparable to that of the stable hybrid aspen-based transgenic lines. Variable expression of the maize A1 gene encoding dihydroflavanol 4-reductase (DFR) as a consequence of imposition of the methylation on the border regions of the integrated single copy of the transgene from the flanking genomic regions has been reported in two transgenic petunia lines (Pröls

and Meyer 1992). Differences in *rolC* expression among the plants of transgenic line Brauna11:35S-*rolC*#2 were observed only under field conditions. Environmental factors may play a key role in the incidence of transgene silencing. Suppression of the maize A1 (DFR) gene was significantly higher in field-grown transgenic petunia plants than in the plants growing in greenhouse and varied with the plant age (Meyer et al. 1992). Similarly, silencing of the chalcone synthase transgene in petunia was observed under high light intensity (van Der Krol et al. 1990), and heat treatment (37 °C) of transgenic tobacco plants led to a reversible reduction or complete loss of transgene-encoded activities in some transformants carrying luciferase and *nptII* genes (Neumann et al. 1997). Transcriptional gene silencing of a transgenic rapeseed line containing a 35S-bialaphos tolerance gene, following infection by cauliflower mosaic virus, has recently been reported (Al-Kaff et al. 2000).

Other interesting observations in the present study were the differences between wild and hybrid aspen clones in their acceptance of the foreign gene and its long-term stability. The differences were observed from the first stage of regeneration and it was not possible to regenerate transgenic plants from wild aspen clones using the low-virulence *Agrobacterium* strain LBA4404 (Fladung et al. 1997). Using this strain, only the hybrid aspen clone was successfully regenerated on kanamycin-containing medium. Although, wild aspen transgenic plants could be obtained using the GV3101 oncogenic *Agrobacterium* strain, a higher in vitro mortality in these transgenic lines was observed compared to the Esch5 hybrid aspen-based transgenic lines. We have so far not observed any in vitro morphological reversion in hybrid aspen-based transgenic lines containing T-DNA repeats. The in vitro cultures of wild aspen-based transgenic lines, however, showed morphological alterations for *rolC* expression, probably as a result of transgene repeats and/or AT-rich flanking regions. The long-term morphological analysis for *rolC* expression also revealed the altered expression in most of the presently available wild aspen-based transgenic lines containing a single copy of the transgene (Table 1). This is in contrast to the Esch5 hybrid aspen-based transgenic lines where several lines are still showing stable *rolC* expression. Therefore, these observations indicate, though do not prove, a specific genomic control over the long-term stable expression of a transgene. The plant factors involved in integration of the T-DNA into the plant genome and its stability have not yet been identified. However, it has been shown recently that histone H2A (*RAT5*) plays an important role in T-DNA integration into the plant genome and the integration step was reported to be blocked in the *RAT5* mutant of *Arabidopsis* (Mysore et al. 2000).

In summary, we have established a molecular basis for transgene silencing in the different aspen lines analyzed in the present study. Similar to previous reports on various plant species (for a review, see Stam et al. 1997a; Kooter et al. 1999), we confirmed that multiple insertions in the form of direct or inverted incomplete/complete T-DNA repeat formation was associated with the loss of transgene expression in the hybrid aspen and in some of the wild aspen transgenic lines. In other wild aspen transgenic lines, however, the transgene inactivation was observed despite single-copy integration, probably as a consequence of AT-rich flanking regions. Therefore, single-copy transgene integration into genomic regions that are free from 'position effect variegation' seems to be a prerequisite for the stable expression of a transgene. In addition, our results indicate that the expression of a single-copy transgene may also be altered even when it is integrated into genomic regions that are non-AT-rich or without repeated DNA. It can be suggested that the genetic composition of the host genome may play a crucial role in transgene integration, and when combined with environmental influences may significantly affect the long-term stable expression of integrated foreign DNA.

Acknowledgements We thank O. Nowitzki, D. Ebbinghaus and A. Schellhorn for technical assistance, and H.-J. Muhs and B. Ziegenhagen [Institut für Forstgenetik und Forstpflanzenzüchtung (BFH), Großhansdorf, Germany] for helpful comments. We also thank anonymous referees for critically reviewing the manuscript. This work is supported by Deutscher Akademischer Austauschdienst, Bonn (to S.K.) and Bundesministerium für Bildung und Forschung, Bonn (to M.F.).

References

Al-Kaff NS, Kreike MM, Covey SN, Pitcher R, Page AM, Dale PJ (2000) Plants rendered herbicide-susceptible by cauliflower mosaic virus-elicited suppression of a 35S promoter-regulated transgene. Nat Biotechnol 18:995–999

Assad FF, Tucker KL, Signer ER (1993) Epigenetic repeat-induced gene silencing (RIGS) in *Arabidopsis*. Plant Mol Biol 22:1067–1085

Cerutti H, Johnson AM, Gillham NW, Boynton JE (1997) Epigenetic silencing of a foreign gene in nuclear transformants of *Chlamydomonas*. Plant Cell 9:925–945

Cluster PD, O'Dell M, Metzlaff M, Flavell RB (1996) Details of T-DNA structural organization from a transgenic petunia population exhibiting co-suppression. Plant Mol Biol 32:1197–1203

Cryderman DE, Morris EJ, Biessmann H, Elgin SCR, Wallrath LL (1999) Silencing of *Drosophila* telomeres: nuclear organization and chromatin structure play critical roles. EMBO J 18:3724–3735

Fladung M (1990) Transformation of diploid and tetraploid potato clones with the *rolC* gene of *Agrobacterium rhizogenes* and characterization of transgenic plants. Plant Breed 104:295–304

Fladung M (1999) Gene stability in transgenic aspen (*Populus*). I. Flanking DNA sequences and T-DNA structure. Mol Gen Genet 260:574–581

Fladung M, Ahuja MR (1995) 'Sandwich' method for non-radioactive hybridizations. BioTechniques 18:3–5

Fladung M, Muhs H-J (1999) Untersuchungen zur Stabilität und Expressivität fremder Gene in Aspenklonen (*Populus tremula* und *P. tremula* × *P. tremuloides*) unter Freilandbedingungen. In: Proc. BMBF-Workshop: Freisetzungsbegleitende Sicherheitsforschung mit genetechnisch veränderten Pflanzen und Mikrorganismen, Braunschweig, 25.05.1998. Druck and Papier Meyer, Scheinfeld, pp 91–99

Fladung M, Muhs H-J, Ahuja MR (1996) Morphological changes observed in transgenic *Populus* carrying the *rolC* gene from *Agrobacterium rhizogenes*. Silvae Genet 45:349–354

Fladung M, Kumar S, Ahuja MR (1997) Genetic transformation of *Populus* genotypes with different chimeric gene constructs: transformation efficiency and molecular analysis. Trans Res 6:111–121

Flavell RB (1994) Inactivation of gene expression in plants as a consequence of specific sequence duplication. Proc Natl Acad Sci USA 91:3490–3496

Fu X, Kohli A, Twyman RM, Christou P (2000) Alternative silencing effects involve distinct types of non-spreading cytosine methylation at a three gene, single-copy transgenic locus in rice. Mol Gen Genet 263:106–118

Gallie DR (1998) Controlling gene expression in transgenics. Curr Opin Plant Biol 1:166–172

Hankeln T, Winterpacht A, Schmidt ER (1996) Instability of tandem repetitive DNA in "natural" and transgenic organisms. In: Schmidt ER, Hankeln T (eds) Transgenic organisms and biosafety. Springer, Berlin Heidelberg New York, pp 181–188

Hobbs SLA, Kpodar P, De Long CMO (1990) The effect of T-DNA copy number, position and methylation on reporter gene expression in tobacco transformants. Plant Mol Biol 15:851–864

Hobbs SLA, Warkentin TD, De Long CMO (1993) Transgene copy number can be positively or negatively associated with transgene expression. Plant Mol Biol 21:17–26

Kilby NJ, Leyser HMO, Furner IJ (1992) Promoter methylation and progressive transgene inactivation in *Arabidopsis*. Plant Mol Biol 20:103–112

Kohli A, Gahakwa D, Vain P, Laurie DA,. Christou P (1999) Transgene expression in rice engineered through particle bombardment: molecular factors controlling stable expression and transgene silencing. Planta 208:88–97

Kooter JM, Matzke, MA Meyer P (1999) Listening to the silent genes: transgene silencing, gene regulation and pathogen control. Trends Plant Sci 4:340–347

Krizkova L, Hrouda M (1998) Direct repeats of T-DNA integrated in tobacco chromosome: characterization of junctions region. Plant J 16:673–680

Kumar S, Fladung M (2000a) Determination of transgene repeat formation and promoter methylation in transgenic plants. BioTechniques 28:1128–1137

Kumar S, Fladung M (2000b) Transgene repeats in aspen: molecular characterization suggests simultaneous integration of independent T-DNAs into receptive hotspots in host genome. Mol Gen Genet 264: 20–28

Linn F, Heidmann I, Saedler H, Meyer P (1990) Epigenetic changes in the expression of the maize A1 gene in *Petunia hybrida*: role of numbers of integrated gene copies and state of methylation. Mol Gen Genet 222:329–336

Matzke AJM, Neuhuber F, Park Y-D, Ambros PF, Matzke MA (1994) Homology-dependent gene silencing in transgenic plants: epistatic silencing loci contain multiple copies of methylated transgenes. Mol Gen Genet 244:219–229

Mette MF, Winden JVD, Matzke MA, Matzke AJM (1999) Production of aberrant promoter transcripts contributes to methylation and silencing of unlinked homologous promoters in trans. EMBO J 18:241–248

Meyer P, Linn F, Heidmann I, Meyer ZAH, Niedenhof I, Saedler H (1992) Endogenous and environmental factors influence 35S promoter methylation of a maize A1 gene construct in transgenic petunia and its color phenotype. Mol Gen Genet 231:345–352

Morino K, Olsen O-A, Shimamoto K (1999) Silencing of an aleurone-specific gene in transgenic rice is caused by rearranged transgene. Plant J 17:275–285

Mysore KK, Nam J, Gelvin SB (2000) An *Arabidopsis* histone H2A mutant is deficient in *Agrobacterium* T-DNA integration. Proc Natl Acad Sci USA 97:948–953

Neumann K, Droge-Laser W, Kohne S, Broer I (1997) Heat treatment results in a loss of transgene-encoded activities in several tobacco lines. Plant Physiol 115:939–947

Park Y-D, Papp I, Moscone EA, 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

Pröls F, Meyer P (1992) The methylation patterns of chromosomal integration regions influence gene activity of transferred DNA in *Petunia hybrida*. Plant J 2:465–475

Salinas J, Matassi G, Montera LM, Bernards G (1988) Compositional compartmentalization and compositional patterns in nuclear genome of plants. Nucleic Acids Res 16:4269–4285

Spena A, Schmülling T, Koncz C, Schell J (1987) Independent and synergistic activity of *rol A*, *B* and *C* loci in stimulating abnormal growth in plants. EMBO J 6:3891–3899

Stam M, Mol JNM, Kooter JM (1997a) The silence of genes in transgenic plants. Ann Bot 79:3–12

Stam M, de Bruin R, Kenter S, van der Hoorn RAL, van Blockland R, Mol JNM, Kooter JM (1997b) Post-transcriptional silencing of chalcone synthase in *Petunia* by inverted transgene repeats. Plant J 12:63–82

Van der Krol AR, Mur LA, Beld M, Mol JN, Stuitje AR (1990) Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. Plant Cell 2:291–299