

Molecular Characterization of a Silenced Transgenic *Nicotiana Tabacum* Line

Antje Dietz-Pfeilstetter (✉ antje.dietz@julius-kuehn.de)

Julius Kuhn-Institut Standort Braunschweig-Messeweg <https://orcid.org/0000-0001-7850-5090>

Ulrike Manske

Julius Kühn-Institut Standort Braunschweig-Messeweg; Julius Kuhn-Institut Standort Braunschweig-Messeweg

Research Article

Keywords: Transgene silencing, *Nicotiana tabacum*, DNA methylation, RNA interference, epigenetics, transcriptional gene silencing

Posted Date: October 27th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-996097/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

RNA interference (RNAi), a natural regulatory mechanism involved in gene regulation, transgene silencing and virus defense in plants, has been widely used for improving the quality of crops and for conferring resistance to plant pests and pathogens. For the sustained application of RNAi-based genetically modified (GM) crops in agriculture, stable expression of the silencing-inducing hairpin transgene is important. As RNA interference may be accompanied by RNA-directed methylation of homologous DNA, including the RNAi transgene, gene body methylation may spread into the promoter and eventually lead to progressive transcriptional silencing of the RNAi inducing transgene. This would be detrimental for the effective use of RNAi crops and may have safety implications if food allergens or antinutrients are the target.

Methods and Results

In order to gain knowledge about the molecular structure of transgenes sensitive to gene silencing as well as about the early detection of unstable plant lines, an unstable GM tobacco line was analyzed with respect to the integration structure and methylation patterns. This plant line was selected because it lost reporter gene expression in progeny plants due to post-transcriptional gene silencing (PTGS) and was subject to transcriptional gene silencing (TGS) during prolonged tissue culture propagation of the primary transformant. It was found to contain an inverted repeat integration with transgene constructs separated by the bacterial transposon Tn5393 and to exist in two epigenetic states with respect to transgene methylation.

Conclusions

Inverted repeat integration containing bacterial transposon sequences may be a possible molecular trigger for the observed silencing of transgene expression in this unstable GM line.

Introduction

Genetically modified (GM) plants with new traits conferred by the expression of transgenes have been used in agriculture in many countries for more than two decades. Despite the rapid evolution and adoption of new plant breeding technologies allowing for targeted genome modifications with increased precision [1], GM plants are still being developed for commercial applications. However, as has been frequently reported for GM plants, gene silencing might unexpectedly occur during the regeneration and testing phase, and sometimes only later during plant development or in transgenic progeny plants [2–5]. Several molecular features are known to act as possible triggers for transgene silencing, such as a suppressive chromatin structure at the integration site, the integration of multiple transgene copies, particularly in inverted-repeat (IR) orientation or transgene rearrangements [6–7].

IR integrations are often found associated with transcriptional gene silencing (TGS) as well as with post-transcriptional gene silencing (PTGS), both mediated by double-stranded (ds) RNA transcribed from the inverted repeat [7–8]. High transgene expression has been identified as another efficient trigger of PTGS through the action of a threshold-dependent RNA sensing mechanism [4, 9]. Since PTGS and TGS are mechanistically linked [10], switches from one silencing state to the other are possible and have been reported to occur during long-term tissue culture and vegetative development of transgenic plants [11–12]. Switches from PTGS to TGS were shown to be correlated with progressive promoter methylation [11, 13].

Post-transcriptional gene silencing triggered by small RNAs (siRNA or microRNA) may not only cause unintended transgene silencing, but is also the basis of RNA interference (RNAi), a natural eukaryotic gene regulatory mechanism. RNAi has been used for the production of GM crops with enhanced food quality traits such as the enrichment with essential amino acids or special fatty acids [14–15]. It has also been exploited to reduce food allergens and toxic metabolites [16–17] and to achieve resistance to plant pathogens and pests [18–21]. Targeted RNAi-mediated gene silencing is typically achieved by the introduction of gene-specific antisense or IR constructs through stable plant transformation, resulting in the expression of silencing-inducing dsRNA or hairpin (hp) RNA. siRNAs derived from these bind to homologous target mRNA, leading to mRNA degradation and inhibition of translation. RNAi acting at the post-transcriptional level may also be associated with RNA-directed DNA methylation (RdDM) of homologous sequences [22] (Jones et al. 1999), and this may include the transcribed region of the silencing construct. If RNAi-induced DNA methylation spreads from transcribed sequences to the upstream transgene promoter due to the production of secondary siRNAs by an RNA-dependent RNA polymerase [23], a switch from PTGS to TGS will occur during plant development, and this will result in an undesirable shut-down of the expression of the silencing construct and phenotypic instability. Although it was found for two endogenous genes that virus-induced gene silencing (VIGS) was not accompanied by RdDM [22, 24] and by target site spread through secondary siRNAs [23], unintended DNA methylation is seen as possible drawback of RNAi applications [25–26]. Loss of silencing has been reported during a field experiment of T₃ generation *Nicotiana attenuata* plants carrying an *acetyl-CoA-transferase 1 (acx1)* silencing construct [13, 27]. If shut-down of dsRNA production occurs, this will not only result in limited stability of the RNAi phenotype over generations [26], but may also raise biosafety issues in the case of RNAi plants with reduced allergens or toxins. With respect to the biosafety assessment of RNAi plants it is therefore important to get more insight into the mechanisms of RNA-directed DNA methylation in transgenic plants.

In previous work on gene expression and molecular analyses of GM *Nicotiana tabacum* plants transformed with a P35S-*gus* gene construct, we identified one line showing a high percentage of gene silencing in F1 progeny as well as progressive gene silencing during long-term tissue culture, accompanied by changes in the DNA methylation state [2]. In order to get more insight into the molecular mechanisms underlying these silencing phenomena, the structure of the T-DNA integration site and the DNA methylation patterns in epigenetic variants were analyzed. In addition, the potential of the silenced transgene locus to direct silencing of homologous active transgene alleles and DNA methylation *in trans* was investigated.

Materials And Methods

Constructs for plant transformation

The binary plasmid vectors pLX222GUS and pBI22SAR2E are based on the binary vector pLX222 [28] and on the *gus* gene fusion plasmid pBI221. They both contain the kanamycin resistance gene *nptII* as selectable marker. The P35S-*gus*-*Tnos* reporter gene constructs present in these plasmids were described in previous work [2]. Binary vectors were mobilized from *Escherichia coli* to the *Agrobacterium tumefaciens* strain LBA4404 for subsequent plant transformation.

Plant transformation and propagation

Transgenic plants were obtained after leaf disk transformation of *N. tabacum* var. W38, followed by shoot and root regeneration, as described previously [2]. Homozygous F3 lines of transformants LX222GUS-A 6b, BI22SAR2E 13a and BI22SAR2E 16b were obtained by two back-crosses of primary transformants with W38 tobacco, followed by selfing. F1 progeny of a young (18 months) LX222GUS-A 9b primary transformant (short TC) were obtained from cuttings. After long-term *in vitro* cultivation (long TC) of transformant LX222GUS-A 9b on MSO agar for 10 years, leaf regenerants were transferred to the greenhouse for the production of 'long TC' T1 progeny by selfing. Seeds were surface sterilized with 0.25 % sodium hypochlorite, washed with distilled water and germinated on MSO agar supplemented with 100 mg/l kanamycin. Segregation ratios for antibiotic resistant and sensitive plants were recorded after 5 weeks. Kanamycin resistant (KanR) transgenic progeny plants were further propagated on MSO medium. Plants were kept in a climate chamber (16 h/8 h light period, 2400 lux) at 22°C, before they were transferred to the greenhouse for further crossings.

Greenhouse cultivation and cross-hybridization

To obtain hybrid progeny, a hemizygous F1 plant (F1-3) and a hemizygous T1 plant (T1-1) of the *sh* or *tTC* respectively the long TC' LX222GUS-A9b epigenetic variant was subject to cross-hybridizations with homozygous F3 plants of three different transgenic *Nicotiana tabacum* lines. In addition backcrosses with non-transgenic *N. tabacum* var. W38 were performed. Young kanamycin resistant seedlings were transferred from the climate chamber to the greenhouse, potted into TKS 1 soil and grown at 24°C (14 h light, 10.000 lux). At flowering time transgenic tobacco plants were emasculated and pollinated using pollen from the different transgenic lines or from non-transgenic *N. tabacum* var. W38. Mature seed pods were harvested and stored at room temperature.

Genomic DNA isolation and molecular analyses

Genomic DNA from plant leaves was extracted either according to [29] or using the NucleoSpin Plant II kit (Macherey-Nagel). Hybrid progeny plants were characterized with respect to the presence or absence of the LX222GUS-A9b transgene insert by Southern blot analysis or by differential polymerase chain reactions (PCR). For Southern blot analysis ten micrograms of plant DNA digested with the restriction enzyme *HindIII* were separated in a 0.8 % agarose gel and transferred to a Hybond-N (Amersham™/ GE Healthcare) nylon membrane [30]. Membrane blots were hybridized according to the manufacturer's protocol with a digoxigenin (DIG)-labeled DNA probe (Roche Applied Science) specific for *nptII* in order to detect left border fragments of individual LX222GUS transgenic lines. For differential PCR the primers 35S-2 (5'-GATAGTGGGATTGTGCG-3') and Npt2 (5'-TAGCCAACGCTATGTCCT-3') amplifying a 1920 base pair (bp) fragment were used to detect the LX222GUS-A9b insert, while GN151 (5'-TCAACTGACACCAAACCTTAG-3') and GN152 (5'-ACACAAACTGCTTATTATCT-3') amplifying a 1096 bp fragment were used as primers for the detection of BI22SAR2E inserts. Amplification reactions containing 0.2 μM primers, 100 μM deoxynucleotides, 1.5 mM MgCl₂, 1x GreenFlexi buffer and 1 unit GoTaq DNA polymerase (both Promega) were incubated after initial denaturation at 94°C in a thermocycler for 35 cycles with annealing temperatures between 52°C and 53°C and 5 min final elongation at 72°C. PCR products were analyzed in 1.8 % agarose gels.

Analysis of the molecular structure of T-DNA integration

For the determination of DNA sequences at left border (LB) and right border (RB) integration sites of LX222GUS-A 9b, inverse PCR assays were performed. T-DNA primers to be used in inverse PCR were pretested by PCRs to ensure the presence of binding sites in the transformed plant DNA. For RB analysis DNA of a 'short TC' plant was cut with the restriction enzyme *AseI* which has a recognition site about 500 bp from the RB, followed by re-ligation. The resulting fragment consisting of known transgene sequences and of unknown flanking sequences was cut with a second restriction enzyme yielding an RB-containing fragment which allowed amplification of flanking DNA sequences by using the T-DNA primers iPCR1 (5'-CATAAAGTGTAAGCCTGG-3') and iPCR2 (5'-GACTCCCTTAATCTCCG-3'). For LB analysis DNA was restricted with *MaeI*, cutting about 490 bp from the putative left border, re-ligated and cut with *StuI*, before flanking DNA was amplified using primers LB1.1 (5'-GTCCGCAATGTGTATTAAAG-3') and iPCR1B (5'-CGATTGCTCATCGTCATT-3'). Cycling conditions for inverse PCRs were: denaturation at 94°C for 2 min; 40 cycles of 94°C for 30 sec, 52°C for 30 sec, 72°C for 2 min + 2 sec/cycle; final elongation at 72°C for 10 min. The PCR products were cloned into a TA vector (pGEM-T; Promega), followed by sequence analysis of individual clones (GATC Biotech/Eurofins Genomics). For further T-DNA integration analysis the following primers were used: Tn3-1 (5'-GATCGGCGGAAAGGTCAACC-3') from the newly identified DNA sequence flanking the right border, iPCR2, lacZ1 (5'-ATGCCTGCAGGTGCTGACTCTA-3'), GUSbi9 (5'-GTGGAGTGAAGAGTATTAGTGTG-3') and the Tn5393-specific primers Tn3-a (5'-GAGCTTCATGGTCTCCAGAA-3'), Tn3-c (5'-TGACCGCTCATTTGGCTCAA-3') and Tn3-d (5'-CATGATGCAGATCGCCATGTA-3') reported by Kim and An [31].

Agroinfiltration assays

Recombinant *A. tumefaciens* strain LBA4404 bacteria transformed with binary vectors pLX222GUS or pBI22SAR2E were grown overnight at 28°C and 200 rpm in 10 ml LB medium supplemented with 100 µg/ml rifampicin and 5 µg/ml tetracycline. Cells were harvested by centrifugation, washed with 20 ml 10 mM 2-[N-morpholino]ethanesulfonic acid (MES), pH 5.7; 10 mM MgCl₂ and re-suspended in MMA medium (10 mM MES, pH 5.7; 10 mM MgCl₂, 200 µM acetosyringone) to an OD₆₀₀ of about 1.0. Subsequently, the cultures were incubated for 2-3 h at room temperature.

Leaves No. 4 to No. 6 from the top of different *Nicotiana tabacum* plants were infiltrated. The bacterial suspensions were applied to the lower side of the leaves using a 2 ml syringe without a needle. Agroinfiltrations were done in replicate for each plant and bacterial strain. Leaf samples were taken before infiltration and from infiltrated areas 6 days post infiltration for further testing.

GUS enzyme assays

β-glucuronidase (GUS) expression was analyzed by a spectrophotometric enzyme assay according to Jefferson [32]. Samples from the sixth leaf from the top were collected from stably transformed plants. 100 mg leaf tissue was homogenized in 100 µl extraction buffer (50 mM sodium phosphate, pH 7.0, 1 mM EDTA, 0.1% Triton X-100). After centrifugation for 5 min at 15 800 g, 4°C the concentration of soluble protein in the supernatant was quantified using the BioRad Protein Assay [33]. For the spectrophotometric GUS assay 35 µg of total protein was used with 1 mM p-nitrophenyl-β-D-glucuronide as substrate. Reactions were stopped at 0, 30 and 60 min by adding 0.4 volumes of 1 M 2-amino-2-methyl-2,3-propanediol. Absorbance of the reaction product p-nitrophenol (p-NP) was measured at 415 nm and calibrated against a dilution series of p-nitrophenol (Sigma).

DNA methylation analysis through bisulfite sequencing

DNA methylation was analyzed through bisulfite conversion of unmethylated cytosine residues and subsequent DNA sequencing. DNA was digested with an appropriate restriction enzyme to allow full denaturation during the conversion step. Bisulfite treatments were carried out using the EZ DNA Methylation-Gold™ Kit (Zymo Research) following the manufacturer's instructions. Bisulfite-treated DNA was subjected to PCR amplifications as described above. Amplification with primers 35Sbi-3 (5'-TAAGGTAAGTAATAGAGATTGGAG-3') and 35Sbi-6 (5'-TCTCTCAAATAAAATAAACTTC-3') resulted in a 584 bp fragment covering the entire P35S promoter sequence as well as vector sequences. With the primers GUSbi-1 (5'-AGGAAGTTATTTATTGGAG-3') and GUSbi-4 (5'-CACCCTTACAAAATCCC-3') a 777 bp fragment was amplified comprising the 5' region of the *gus* coding sequence. For bisulfite methylation analysis of the 3' region of the *gus* coding sequence and the adjacent *Tnos* terminator, the primers GUSbi9 (5'-GTGGAGTGAAGAGTATTAGTGTG-3') and LXbi1 (5'-ACCTCTTCACTATTACRCCRC-3') were used to amplify a 692 bp fragment specific for LX222GUS-A inserts, while GUSbi9 (5'-GTGGAGTGAAGAGTATTAGTGTG-3') and SARbi1 (5'-AATACTCCCACTAATCATAATTC-3') were used to amplify a 719 bp fragment specific for the BI22SAR2E inserts. The PCR products were cloned into pGEM-T (Promega) and 2 to 10 clones from each amplification reaction were sequenced (GATC Biotech/Eurofins Genomics). Sequence data were analyzed and methylation densities were calculated using CyMATE software [34]. As is shown in Fig. 2a, CyMATE analysis of the 35S promoter started at position -343 of the promoter sequence and ended at the TATA box (-24). For the 5' *gus* gene, analysis started 43 bp upstream of the start codon and continued up to position 693 of the coding sequence. For the 3' *gus-Tnos* region it started at position 1583 of the coding sequence, encompassed the complete *nos*-terminator and ended about 95 bp respectively 120 bp downstream.

RNA extraction and siRNA detection by Northern blot analysis

RNA from plant leaves was extracted using NucleoZOL (Macherey-Nagel), thereby separating small RNAs (10 – 200 nt) from the large RNA fraction (> 200 nt). For Northern blot analysis 2 µg of denatured small RNAs were separated overnight on a 15% denaturing polyacrylamide/7 M urea gel and transferred onto a positively charged nylon membrane through electroblotting (Semi-dry Blotter/Hoefer). Following chemical crosslinking with 1-Ethyl-3-(3-dimethylaminopropyl)-Carbodiimid (EDC) [35] the membrane was hybridized with digoxigenin (DIG)-labeled RNA probes (Roche Applied Science) specific for the 3'-end (660 bp fragment) of the *gus* gene. Hybridizations were conducted overnight at 40°C in DIG Easy Hyb (Roche). The membrane was washed twice with 2x saline-sodium citrate (SSC); 0.1% SDS for 10 min at 40°C and then twice with 0.1x SSC; 0.1 % SDS for 15 min at 40°C. Immunological detection was performed with 1:10,000 anti-DIG-alkaline phosphatase (AP) (Roche). For AP detection the chemiluminescent substrate CDP-Star (Roche/Sigma-Aldrich) was used, followed by signal exposure on X-ray film (Agfa).

Results

Switch from post-transcriptional gene silencing (PTGS) to transcriptional gene silencing (TGS) during prolonged tissue culture propagation

As previously reported [2], about 60% of the transgenic progeny plants obtained from the transformant LX222GUS-A9b after short-term tissue culture (18 months) lost expression of the *gus* gene, while 40 % showed reduced expression under tissue culture conditions. When the LX222GUS-A9b primary transformant was vegetatively propagated over a period of 10 years, GUS expression was completely and irreversibly shut off, and this complete silencing was also found in all of the T1 progeny plants. In order to confirm the presumed switch from incomplete PTGS-based gene silencing to TGS we investigated the potential for *trans*-silencing of homologous transgenes. *Trans*-silencing mediated by small interfering RNA (siRNA) has been reported by Fojtova et al. [36] for a post-transcriptionally silenced transgene locus, but not for the respective TGS variant. We therefore introduced homologous transgene constructs by agroinfiltration into plants of both tissue culture (TC) variants. As shown in Fig. 1a, in the silenced 'short TC' plant F1-3 there was no expression from two

different agroinfiltrated homologous P35S-*gus*-Tnos gene constructs, confirming PTGS *in trans*. In contrast, no *trans*-effect on *gus* gene expression was found for the 'long TC' variant, similar to non-transformed W38 tobacco plants.

In order to investigate the influence of *sh* or *tTC'* and long TC' LX222GUS-A 9b variants on the expression of stably integrated homologous transgenes, cross-hybridizations of hemizygous plants with different non-silenced transgenic *N. tabacum* lines were carried out. As expected, there was a clear correlation between the presence of the 'short TC' LX222GUS-A9b epiallele and silencing of the *gus* gene (Fig. 1b). Silencing was found to be also correlated with the detection of siRNA homologous to the 3'-end of the *gus* gene (Fig. 1c). Unexpectedly and in contrast to the agroinfiltration experiments, the presence of the LX222GUS-A9b 'long TC' epiallele affected the expression of a homologous transgene in hybrids. Following a temporary suppression, a general increase in expression usually occurred after transfer of cuttings to the greenhouse (Fig. 1d).

Switch from PTGS to TGS accompanied by progressive DNA methylation

It has been previously shown that the 35S promoter and the 5' *gus* region of silenced 'short TC' LX222GUS-A 9b F1 plants were not methylated [2]. However, upon further analysis, we found the 3' *gus*-Tnos region of a silenced (F1-3) as well as of a GUS expressing (F1-1) F1 plant to be strongly methylated, in contrast to plants of stable GUS-expressing GM lines LX222GUS-A 6b and BI22SAR2E 16b (Fig. 2b). The DNA methylation status was inherited to the analyzed LX222GUS-A 9b F2 plant derived from F1-3. Trans-generational spread of methylation into the 5'-part of the *gus* gene was not observed for this epigenetic variant (data not shown). In contrast to 'short TC' plants and in accordance with the observed switch to transcriptional silencing, strong *gus* gene methylation and varying levels of methylation of the 35S promoter were detected in 'long TC' T₀ plants, and these epigenetic DNA modifications were conserved in a leaf regenerant and inherited to its T1 progeny plants (Fig. 2c-d).

Transposon sequence identified at T-DNA integration site in line LX222GUS-A 9b

The progressively silenced transformant LX222GUS-A 9b contains two T-DNA copies that are tightly linked, since they could not be separated by outcrossing. However, no direct tandem or inverted repeat integration could be detected, when using conventional PCR analysis [2]. We therefore aimed to identify sequences at the integration sites to get clues on molecular features that may possibly cause unstable expression and silencing of the *gus* gene. Inverse PCR yielded a right border (RB) integration site fragment that had no sequence homology to the *N. tabacum* genome. However, BLASTn analysis revealed a 566 bp sequence adjacent to the right border that has 99% identity with Ti plasmid sequences in *Agrobacterium tumefaciens* strain LBA4404 encoding part of a Tn3 family transposase (Sequence ID: KY000037.1). It was previously shown that *A. tumefaciens* strain LBA4404 carries the Tn3 family transposon Tn5393 which was found to be introduced into the T-DNA of binary vectors, resulting in the presence of Tn5393 sequences in transgenic rice plants [31]. The 3' end of the transposase gene (*tnpA*) located on the transposable element Tn5393, including the *tnpA* stop codon as well as the 81 bp 5'-terminal inverted repeat (5'-IR; IR-L) of Tn5393, was identified within the 566 bp sequence. PCR analysis using specific primers from Tn5393 revealed that the transposon was completely integrated close to the right border of the second T-DNA copy. No second RB fragment could be identified by inverse PCR, however a 21 bp inverted repeat of RB sequences separated by 7 nucleotides was created during plant transformation. Adjacent to the 21 bp RB sequences of the second T-DNA is the entire Tn5393, terminated by the 81 bp 3'-terminal inverted repeat (3'-IR; IR-R), followed by continuing sequences from the right border of the second T-DNA. Figure 3 shows the molecular structure of the T-DNA insertion as well as details of the DNA sequences at the borders of Tn5393 and integrated T-DNA copies.

DNA methylation in hybrids of LX222GUS-A 9b

It was investigated whether siRNA-directed DNA methylation occurred *in trans* at homologous loci in hybrid plants containing the LX222GUS-A9b 'short TC' epiallele. Due to sequence differences between LX222GUS-A and BI22SAR2E transgenic plants in the region downstream of the *nos* terminator, differential methylation analysis was possible for hybrids obtained from crossings with BI22SAR2E 13a and BI22SAR2E 16b. As is shown in Fig. 4, the 3' *gus*-Tnos region remained methylated at the LX222GUS-A9b 'short TC' epiallele, mainly at symmetrical C positions, while no or very little DNA methylation was found for the BI22SAR2E transgene loci.

Discussion

Shut-down of GUS reporter gene expression during prolonged tissue culture accompanied by changes in DNA methylation

Epigenetic changes within the transgene construct of the GM plant LX222GUS-A9b which is progressively silenced during tissue culture propagation have been investigated. Methylation predominantly at symmetrical C positions (Fig. 2b) that was found for the 'short TC' epigenetic variant in the most distal part of the *gus* gene (starting 230 bp upstream of the stop codon) and within the *nos* terminator suggests DNA methylation mechanisms, which at least partly are independent of siRNA. As is stated in a review by Matzke et al. [37], once established, methylation at symmetrical CG and CHG sites is copied during DNA replication by maintenance methyltransferases. One of these, chromomethylase 3 (CMT3) which catalyzes CHG methylation, interacts with repressive histone modifications [38-39]. In line with this, Dalakouras et al. [40] reported that in stably integrated transgenes CG methylation was most prominent, while it has been proposed that establishment and maintenance of CHG methylation probably depend on structural features of the transgene insertion [41-42]. Phillips et al. [43] discovered recently in *Nicotiana benthamiana* that progressive T-DNA methylation at promoter and gene body sequences in transiently agroinfiltrated leaves occurred primarily in the CHH context and less at CHG and CG sites. These authors suggest that early T-DNA methylation prior to integration is one of the factors responsible for variable and unstable expression in GM plants. For LX222GUS-A9b there is no evidence for early T-DNA methylation at P35S and 5' *gus* gene sequences, however early methylation may have occurred in T-DNA sequences downstream of the *gus* gene. The methylation status of the 3' *gus*-

Tnos region in the young primary LX222GUS-A9b transformant, which shows stable GUS expression during development, has not been determined. However, in the 'short TC' progeny plant F1-1 GUS expression is not silenced despite methylation in the 3' *gus-Tnos* region, and it is therefore unlikely that this methylation was triggered by siRNAs homologous to the 3'-part of the *gus* gene. It is also interesting to note that in plant F1-1 methylation in the 3' *gus-Tnos* region is largely restricted to CG sites. Methylation confined to CG contexts can also be found within active genes and has been shown to be independent of RNA and to be mediated by maintenance methyltransferase MET1 [44]. In contrast to the young LX222GUS-A9b transformant and to F1-1, plant F1-3 expresses *gus* at a very low level, and this silencing is probably due to siRNA homologous to the 3'-part of the *gus* gene which may also trigger the initiation of additional DNA methylation in the 3' *gus-Tnos* region in CG, CHG and CHH sequence contexts. C-methylation in all sequence contexts is typical for RdDM [44-45]. PTGS through *gus*-specific siRNAs in F1-3 is confirmed by the observed *trans*-silencing of homologous *gus* gene constructs (Fig. 1a-c).

Through prolonged tissue culture DNA methylation spread upstream into the 5'-part of the *gus* coding sequence and into the 35S promoter, causing a total shut-down of GUS expression. Progressive promoter methylation and silencing during prolonged *in vitro* cultivation or vegetative propagation of GM plants has been previously observed in potato and tobacco [11-12]. In the 'long TC' LX222GUS-A 9b primary transformant all analyzed copies of the 5'-*gus* gene region were extensively methylated at symmetrical and non-symmetrical C positions, and this methylation pattern was transmitted to a leaf regenerant and its T1 progeny plants (Fig. 2c). Weaker and more variable methylation was detected in the 35S promoter of the 'long TC' primary transformant and of its descendants obtained from secondary regeneration and sexual propagation (Fig. 2d). A spread of methylation from the 3'-part of a transcribed gene region into the 5'-gene region and the 35S promoter was reported in the study of Fojtova et al. [11] where a post-transcriptionally silenced *N. tabacum* line containing a T-DNA inverted repeat gave rise to an epigenetic variant after 24 months of *in vitro* cultivation. In the present study *de novo* methylation of 5'-*gus* sequences may have been caused by primary and secondary siRNAs, produced from aberrant transgene RNAs (particularly transcripts lacking 3'-polyadenylation) by RNA-dependent RNA polymerase 6 (RDR6) during a temporary PTGS phase, similar to what has been suggested by Bond and Baulcombe [46] for the new establishment of transposable element methylation. Aberrant transcripts have been shown to be strong inducers of PTGS of transgenes in plants [47]. Alternatively, as suggested by Fojtova et al. [12], spread of methylation may have been triggered by chromatin modification. Zemach et al. [39] found that in *Arabidopsis thaliana* most non-symmetric methylation of plant transposable elements is mediated by chromomethylase CMT2 and is facilitated by the nucleosome remodeler DDM1 independent of RdDM. The initial chromatin and methylation status of the integrated transgene DNA, probably determined by specific features at the integration site, might have affected the progression of silencing observed in GM tobacco line LX222GUS-A 9b.

While GUS expression was silenced in progeny plants of the 'short TC' LX222GUS-A 9b transformant, silencing of the selectable marker gene *nptII* in these plants was not observed. In contrast, for 'long TC' progeny the ratio of kanamycin resistant and sensitive plants was 30:17 and deviated from the expected 3:1 ratio of progeny containing and not-containing the *nptII* gene, that was confirmed by PCR analysis. It is therefore likely that silencing eventually proceeds beyond the *gus* gene. In line with this, we found evidence for enhanced methylation of *Pnos* regulatory sequences in the 'long TC' primary transformant (data not shown).

Bacterial transposon as possible target and trigger for PTGS and TGS

In the present study, the bacterial *tnpA* gene as well as the complete bacterial transposon Tn5393 were identified in the RB flanking region of the first of two LX222GUS-A 9b T-DNA copies (Fig. 3) integrated into the plant genome. Tn5393 most likely originated from the *A. tumefaciens* strain LBA4404 that was used for plant transformation [31]. From PCR and DNA sequence analysis it is evident that Tn5393 was initially present within the second T-DNA, close to its right border. Upon inverted repeat T-DNA integration Tn5393 sequences were placed between right border sequences of the two T-DNA copies. There is no indication for additional transposon sequences integrated within the first T-DNA copy. This is in line with findings of Kim and An [31] who often detected normal T-DNA and Tn5393-carrying T-DNA in the same transgenic plant. These authors also found that in about 26% of 331 randomly selected transgenic rice lines transformed with binary vectors using LBA4404 Tn5393 was co-integrated. No bacterial transposon sequences were detected in the other GM lines in our study using different PCR primer combinations from Tn5393. When analyzing the widely used *Nicotiana benthamiana* line 16c, Philips et al. [48] found the 5' IR, the complete *tnpA* gene, the recombination region (*res*) and part of the resolvase (*tnpR*) gene of Tn5393 co-integrated with the T-DNA. These authors speculate that the bacterial transposon sequence adjacent to the reporter transgene may have an enhancing effect on the plant's response to silencing inducers, making that line a sensitive tool for epigenetics research. Eukaryotic transposons are often targets for DNA methylation and silencing [49-51], and similarly bacterial transposon sequences transferred to plants could be nucleation centers for DNA methylation. It is therefore conceivable that in the case of the GM plant line LX222GUS-A 9b the Tn5393 sequences adjacent to the right borders of the T-DNA copies had an effect on the chromatin structure and on the methylation status of the integrated T-DNA. Unlike *N. benthamiana* line 16c which is a non-silenced line expressing green fluorescent protein (GFP), GM tobacco line LX222GUS-A 9b contains two T-DNA copies and undergoes progressive *gus* gene silencing without an external silencing inducer. Silencing, however, does not occur immediately upon transgene integration, but only in the next generation or during extended tissue culture. The exact trigger for siRNA production and initiation of PTGS in this line is not known. It is likely that Tn5393 with its terminal IR sequences in combination with the inverted repeat T-DNA integration structure was involved in siRNA production and establishment of DNA methylation and presumably had an enhancing effect on the progression of silencing, as has been suggested by Philips et al. [47].

Effects on the expression and methylation of homologous transgenes

In the agroinfiltration assays two different homologous transgene constructs introduced for transient expression of the *gus* gene were subject to gene silencing in a non-expressing 'short TC' F1 plant, confirming our suspicion that this variant conferred *trans*-PTGS via small interfering RNA (siRNA), similar to the *trans*-silencing described by Fojtova et al. [36]. The suppressive effect of the 'short TC' epigenetic variant was also found in hybrids carrying the LX222GUS-A 9b F1-3 locus obtained after cross-hybridization of three GM lines with homologous transgene sequences. The fact that in two of these lines the P35S-*gus-Tnos* gene construct was flanked by a scaffold/matrix attachment region had no effect on the *trans*-silencing efficiency of the silencing-triggering

locus. As is shown in Fig. 1b-c, there was a strict correlation between presence of the LX222GUS-A 9b epiallele, gene silencing, and the detection of siRNA. These observations confirm the non-heritable nature of PTGS in the absence of the silencing triggering locus. PTGS relies on siRNA as trigger for mRNA decay or translational inhibition, and this signal has to be produced via dsRNA from a silencing-inducing locus.

When the 'long TC' epigenetic variant was used as recipient for agroinfiltration of homologous transgene constructs, no suppressive effect on the transient expression of the newly introduced *gus* gene was found, in accordance with the transcriptional silencing state of this epiallele. However, in hybrid progeny plants of this epigenetic variant variable *trans* effects on gene expression were observed (Fig. 1c), and this may be due to low levels of siRNA homologous to the *gus* gene which were detected in these hybrids (data not shown). 30-50% reduced expression of homologous transgenes conferred by a TGS epiallele has also been reported by Fojtova et al. [36].

Despite the presence of siRNA, mainly methylation at symmetrical CG and CHG sites of the 3' *gus-Tnos* region was detected in progeny plants derived from the 'short TC' plant F1-3 (Fig. 2b and Fig. 4). This methylation pattern confirms the presumed predominance of maintenance methylation by MET1 and CMT3 in this epigenetic variant and the assumption that methylation in the 3' *gus-Tnos* region is not primarily mediated by siRNAs. In accordance with this, no or only negligible DNA methylation was found in the 3' *gus-Tnos* region of homologous transgene loci in hybrids. In contrast to these findings, Fojtova et al. [36] observed that the PTGS variant of an inverted repeat transgene locus did not only confer silencing *in trans* to homologous transgene loci, but conferred *trans*-methylation at both symmetrical and non-symmetrical C positions as well. As already mentioned above, the specific integration structure encompassing a bacterial Tn3 transposon and the chromatin status of GM line LX222GUS-A 9b may have been responsible for the initiation of DNA methylation independent of RdDM, while methylation at symmetrical C sites is then transmitted to successive generations during DNA replication by the action of maintenance methyltransferases.

Statements And Declarations

Compliance with ethical standards

Potential conflict of interests

The authors have no relevant financial or non-financial interests to disclose.

Funding

This study was supported by the Federal Ministry of Food and Agriculture, but no special funding received.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Author's contributions

Both authors contributed to the design and performance of the experiments. Data collection and analyses were performed by ADP, who also wrote the original draft. Both authors read and approved the final manuscript.

Data availability

DNA sequencing data generated and analyzed during the current study are available in the GenBank Database.

References

1. Modrzejewski D, Hartung F, Sprink T, Krause D, Kohl C, Wilhelm R (2019) What is the available evidence for the range of applications of genome-editing as a new tool for plant trait modification and the potential occurrence of associated off-target effects: a systematic map. *Environ Evid* 8:27. DOI:10.1186/s13750-019-0171-5
2. Dietz-Pfeilstetter A, Arndt N, Manske U (2016) Effects of a petunia scaffold/matrix attachment region on copy number dependency and stability of transgene expression in *Nicotiana tabacum*. *Transgenic Res* 25:149–162
3. De Wilde C, Podevin N, Windels P, Depicker A (2001) Silencing of antibody genes in plants with single-copy transgene inserts as a result of gene dosage effects. *Mol Genet Genomics* 265:647–653
4. Elmayan T, Vaucheret H (1996) Expression of single copies of a strongly expressed 35S transgene can be silenced post-transcriptionally. *Plant J* 9:787–797
5. Brandle JE, McHugh SG, James L, Labbé H, Miki BL (1995) Instability of transgene expression in field grown tobacco carrying the *csr1-1* gene for sulfonyleurea herbicide resistance. *Biotechnol* 13:994–998

6. Morel JB, Mourrain P, Beclin C, Vaucheret H (2000) DNA methylation and chromatin structure affect transcriptional and posttranscriptional transgene silencing in Arabidopsis. *Current Biol* 10:1591–1594
7. De Buck S, Van Montagu M, Depicker A (2001) Transgene silencing of invertedly repeated transgenes is released upon deletion of one of the transgenes involved. *Plant Mol Biol* 46:433–445
8. Muskens MWM, Vissers APA, Mol JNM, Kooter JM (2000) Role of inverted DNA repeats in transcriptional and post-transcriptional gene silencing. *Plant Mol Biol* 43:243–260
9. Schubert D, Lechtenberg B, Forsbach A, Gils M, Bahadur S, Schmidt R (2004) Silencing in Arabidopsis T-DNA transformants: The predominant role of a gene-specific RNA sensing mechanisms versus position effects. *Plant Cell* 16:2561–2572
10. Sijen T, Vijn I, Rebocho A, Van Blokland R, Roelofs D, Mol JN, Kooter JM (2001) Transcriptional and posttranscriptional gene silencing are mechanistically related. *Curr Biol* 11:436–440
11. Fojtová M, Van Houdt H, Depicker A, Kovařík A (2003) Epigenetic switch from posttranscriptional to transcriptional silencing is correlated with promoter hypermethylation. *Plant Physiol* 133:1240–1250
12. Nocarova E, Opatrný Z, Fischer L (2010) Successive silencing of tandem reporter genes in potato (*Solanum tuberosum*) over 5 years of vegetative propagation. *Ann Bot* 106:565–572
13. Weinhold A, Kallenbach M, Baldwin T (2013) Progressive 35S promoter methylation increases rapidly during vegetative development in transgenic *Nicotiana attenuata* plants. *BMC Plant Biol* 13:99. Doi:10.1186/1471-2229-13-99
14. Liu Q, Wu M, Zhang B, Shrestha P, Petrie J, Green AG, Singh SP (2016) Genetic enhancement of palmitic acid accumulation in cotton seed oil through RNAi down-regulation of ghKAS2 encoding β -ketoacyl-ACP synthase II (KASII). *Plant Biotech J*. Doi:10.1111/pbi.12598
15. Houmar NM, Mainville JL, Bonin CP, Huang S, Luethy MH, Malvar TM (2007) High-lysine corn generated by endosperm-specific suppression of lysine catabolism using RNAi. *Plant Biotech J* 5:605–614
16. Dodo HW, Konan KN, Chen FC, Egnin M, Viquez OM (2008) Alleviating peanut allergy using genetic engineering: the silencing of the immunodominant allergen Ara h 2 leads to its significant reduction and a decrease in peanut allergenicity. *Plant Biotech J* 6:135–145
17. Sunilkumar G, Campbell LM, Puckhaber L, Stipanovic RD, Rathore KS (2006) Engineering cottonseed for use in human nutrition by tissue-specific reduction of toxic gossypol. *Proc Natl Acad Sci USA* 103:18054–18059
18. McLoughlin AG, Walker PL, Wytinck N, Sullivan DS, Whyard S, Belmonte MF (2018) Developing new RNA interference technologies to control fungal pathogens. *Can J Plant Pathology* 40:325–335
19. Ibrahim AB, Aragão FJ (2015) RNAi-mediated resistance to viruses in genetically engineered plants. *Methods Mol Biol* 1287:81–92
20. Koch A, Kogel K (2014) New wind in the sails: improving the agronomic value of crop plants through RNAi-mediated gene silencing. *Plant Biotech J* 12:821–831
21. Rosa C, Kuo Y-W, Wuriyangan H, Falk BW (2018) RNA interference mechanisms and applications in plant pathology. *Ann Rev Phytopathology* 56:581–610
22. Jones L, Hamilton AJ, Voinnet O, Thomas CL, Maule AJ, Baulcombe DC (1999) RNA-DNA interactions and DNA methylation in post-transcriptional gene silencing. *Plant Cell* 11:2291–2301
23. Vaistij FE, Jones L, Baulcombe DC (2002) Spreading of RNA targeting and DNA methylation in RNA silencing requires transcription of the target gene and a putative RNA-dependent RNA polymerase. *Plant Cell* 14:857–867
24. Thomas CL, Jones L, Baulcombe DC, Maule AJ (2001) Size constraints for targeting post-transcriptional gene silencing and for RNA-directed methylation in *Nicotiana benthamiana* using a potato virus X vector. *Plant J* 25:417–425
25. Dalakouras A, Papadopoulou KK (2020) Epigenetic modifications: an unexplored facet of exogenous RNA applications in plants. *Plants* 9:673. Doi:10.3390/plants9060673
26. Casacuberta JM, Devos Y, du Jardin P, Ramon M, Vaucheret H, Nogué F (2015) Biotechnological uses of RNAi in plants: risk assessment considerations. *Trends Biotech* 33:145–147
27. Kallenbach M, Bonaventure G, Gilardoni PA, Wissgott A, Baldwin IT (2012) Empoasca leafhoppers attack wild tobacco plants in a jasmonate-dependent manner and identify jasmonate mutants in natural populations. *Proc Natl Acad Sci USA* 109:1548–1557
28. Landsmann J, Llewellyn D, Dennis ES, Peacock WJ (1988) Organ regulated expression of the *Parasponia andersonii* haemoglobin gene in transgenic tobacco plants. *Mol Gen Genet* 214:68–73
29. Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA miniprep: version II. *Plant Mol Biol Rep* 1:19–21
30. Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98:503–517
31. Kim S-R, An G (2012) Bacterial transposons are co-transferred with T-DNA to rice chromosomes during *Agrobacterium*-mediated transformation. *Mol Cells* 33:583–589
32. Jefferson RA (1987) Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Mol Biol Rep* 5:387–405
33. Bradford MM (1976) A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254
34. Hetzl J, Foerster AM, Raidl G, Mittelsten Scheid O (2007) CyMATE: a new tool for methylation analysis of plant genomic DNA after bisulfite sequencing. *Plant J* 51:526–536
35. Pall GS, Hamilton AJ (2008) Improved northern blot method for enhanced detection of small RNA. *Nat Protoc* 3:1077–1084

36. Fojtová M, Bleys A, Bedřichová J, Van Houdt H, Křížová K, Depicker A, Kovařík A (2006) The *trans*-silencing capacity of invertedly repeated transgenes depends on their epigenetic state in tobacco. *Nucleic Acids Res* 34:2280–2293
37. Matzke MA, Kanno T, Matzke AJM (2015) RNA-directed DNA methylation: the evolution of a complex epigenetic pathway in flowering plants. *Ann Rev Plant Biol* 66:243–267
38. Bond DM, Baulcombe DC (2014) Small RNAs and heritable epigenetic variation in plants. *Trends Cell Biol* 24:100–107
39. Zemach A, Kim MY, Hsieh P-H, Coleman-Derr D, Eshed-Williams L, Thao K et al (2013) The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* 153:193–205
40. Dalakouras A, Dadami E, Zwiebel M, Krczal G, Wassenegger M (2012) Transgenerational maintenance of transgene body CG but not CHG and CHH methylation. *Epigenetics* 7:1071–1078. <http://dx.doi.org/10.4161/epi.21644>
41. Dalakouras A, Tzanopoulou M, Tsagris M, Wassenegger M, Kalantidis K (2011) Hairpin transcription does not necessarily lead to efficient triggering of the RNAi pathway. *Transgenic Res* 20:293–304
42. Aufsatz W, Mette MF, van der Winden J, Matzke AJ, Matzke M (2002) RNA-directed DNA methylation in Arabidopsis. *Proc Natl Acad Sci USA* 99:16499–16506
43. Philips JG, Dudley KJ, Waterhouse PM, Hellens RP (2019) The rapid methylation of T-DNAs upon *Agrobacterium* inoculation in plant leaves. *Front Plant Sci* 10:312. Doi:10.3389/fpls.2019.00312
44. Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* 11:204–220
45. Pélissier T, Thalmeir S, Kempe D, Sängler HL, 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
46. Bond DM, Baulcombe DC (2015) Epigenetic transitions leading to heritable, RNA-mediated *de novo* silencing in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 112:917–922
47. Luo Z, Chen Z (2007) Improperly terminated, unpolyadenylated mRNA of sense transgenes is targeted by RDR6-mediated RNA silencing in *Arabidopsis*. *Plant Cell* 19:943–958
48. Philips JG, Naim F, Lorenc MT, Dudley KJ, Hellens RP, Waterhouse PM (2017) The widely used *Nicotiana benthamiana* 16c line has an unusual T-DNA integration pattern including a transposon sequence. *PLoS ONE* 12(2):e0171311. Doi:10.1371/journal.pone.0171311
49. Nuthikattu S, McCue AD, Panda K, Fultz D, DeFraia C, Thomas EN et al (2013) The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21-22 nucleotide small interfering RNAs. *Plant Physiol* 162:116–131
50. Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. *Nature Rev Genetics* 8:272–285
51. Arnaud P, Goubely C, Pélissier T, Deragon J-M (2000) SINE retroposons can be used *in vivo* as nucleation centers for *de novo* methylation. *Mol Cell Biol* 20:3434–3441

Figures

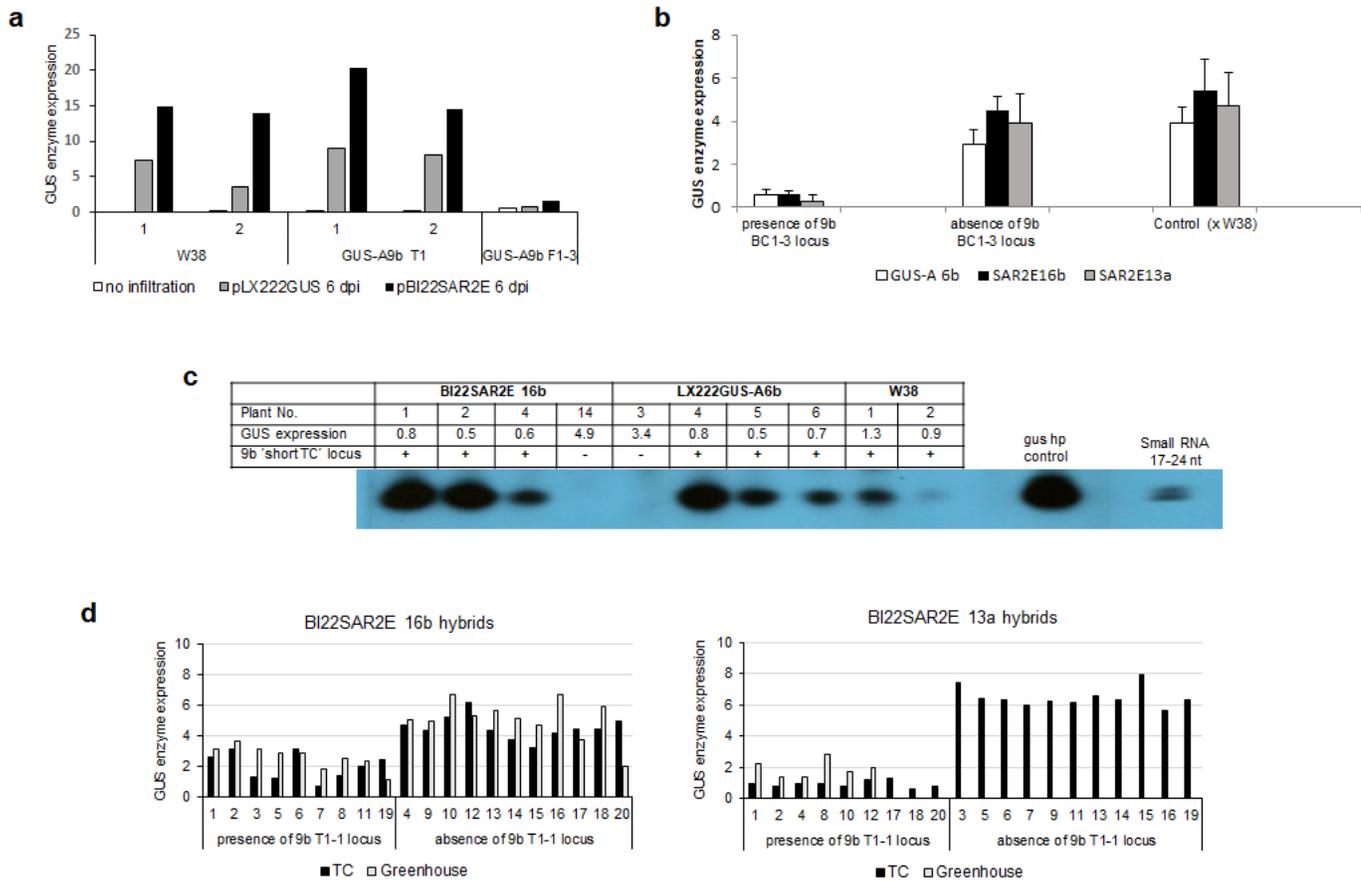


Figure 1

GUS expression given as enzyme activity (nmol p-nitrophenol per mg protein and per minute) a in leaves of individual first generation progeny plants of LX222GUS-A 9b *longTC'* (*T1 - 1* and *T1 - 2*) and short *TC'* (*F1-3*) variants before and 6 days after agroinfiltration of two transgene constructs homologous to the integrated transgene. Two non-transgenic *N. tabacum* (*W38*) plants were used as control. Bars represent mean values of two repeat experiments, b in leaves of hybrids containing (presence) or not containing (absence) the epiallele of the LX222GUS-A 9b *sh* or *tTC'* plant *F1 - 3*. Controls show the rest of or backcross progeny of *esLX222GUS - A6b*, *BI22SAR2E16b* and *BI22SAR2E13a*. c short *TC'* epiallele and detection of siRNA in individual hybrid and backcross plants. siRNA produced in a *gus*-hairpin (*hp*) plant as well as specific 17-24 nt oligonucleotides served as positive controls for siRNA detection by hybridization with a 3'*gus*-specific RNA probe, d GUS expression in leaves of hybrids grown in tissue culture (TC) or in the greenhouse which contain (presence) or do not contain (absence) the 'long *TC'* epiallele from plant *T1-1*. For BI22SAR2E 13a hybrids GUS expression in the greenhouse was analyzed only in plants No. 1, 2, 4, 8, 10 and 12

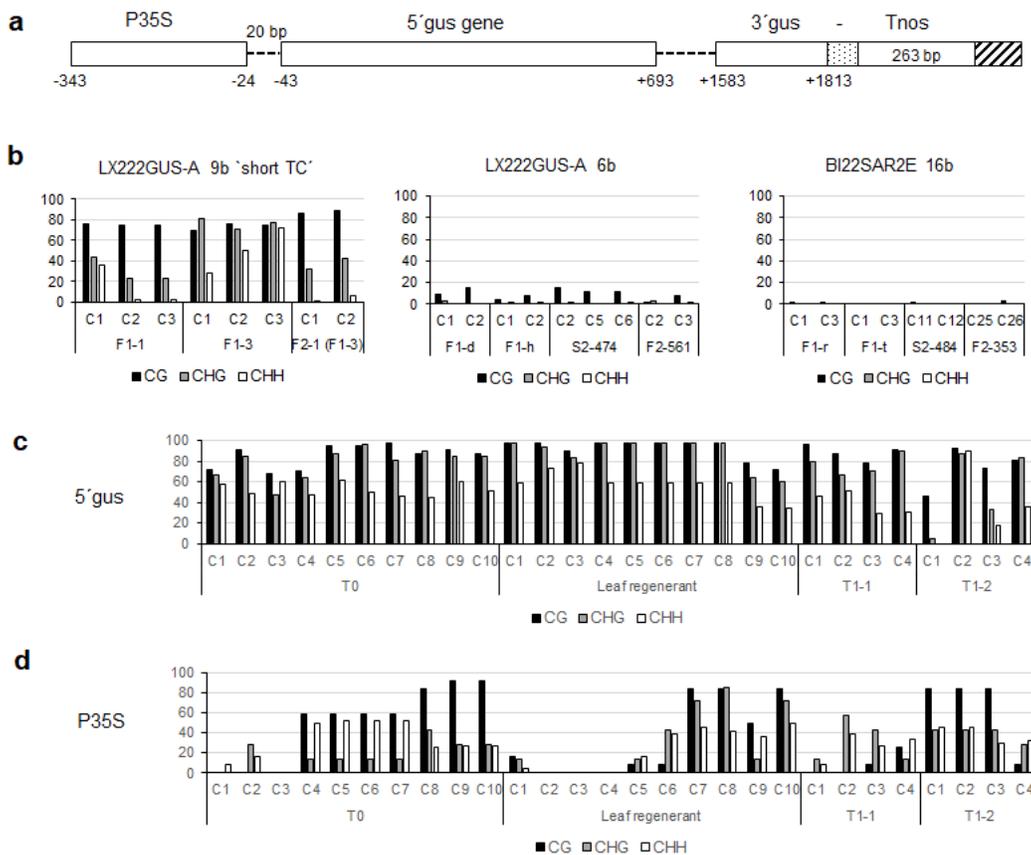


Figure 2

DNA methylation percentages at CG, CHG and CHH sites in successive generations of LX222GUS-A 9b tissue culture (TC) variants. Data of different clones derived from individual plants are shown. a Schematic outline of the genetic regions analyzed. Dots delineate the 70 bp region between the gus 3' end and Tnos while the striped area shows the region deviating between LX222GUS and BI22SAR2E transformants, b Analysis of 3'gus-Tnos region in F1 and F2 plants of the LX222GUS-A 9b *sh* or *tTC'* variant and *ofnon-si* \leq *ncedtrans* \geq *nicl* \in *es*, *c-dAnalysisofthe5'gus* \geq \neq *region(736bp)* and *ofP35S(319bp, c or epromoter)* long TC' variant, in a secondary leaf regenerant and in T1 progeny plants

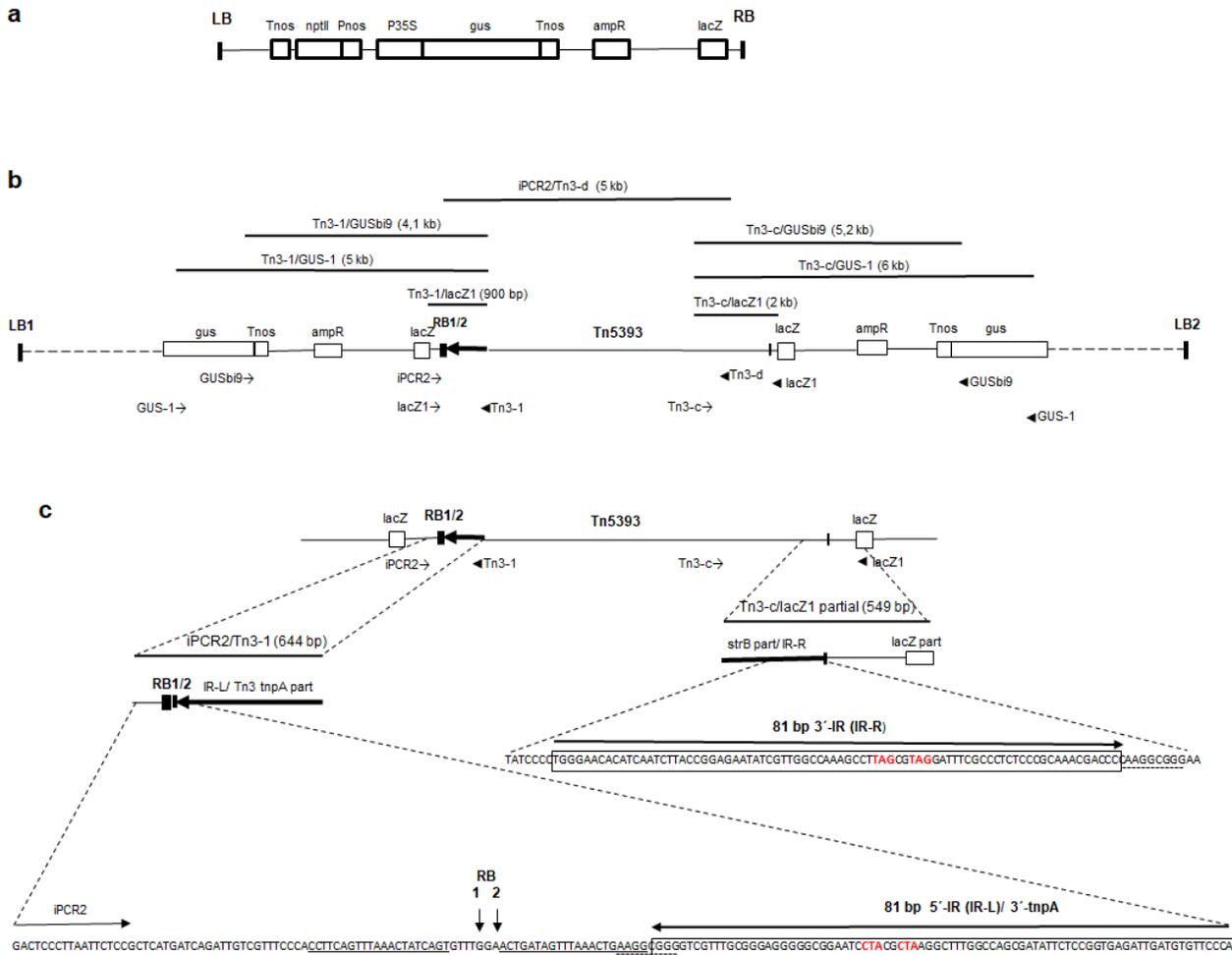


Figure 3

a Genetic elements within the T-DNA of the binary vector pLX222GUS, b T-DNA integration structure of LX222GUS-A 9b as identified by inverse PCR, DNA sequencing and PCR analysis using different primer combinations. Lines above represent PCR products. LB1/ RB1 and LB2/ RB2 are left and right borders of first and second T-DNA copies, respectively, c Sequence analysis of PCR fragments from Tn5393/ pLX222GUS junction regions (644 bp of iPCR2/Tn3-1 fragment, GenBank accession number MW494967; and 549 bp of Tn3-c/lacZ1 fragment, GenBank accession number MW494968; as indicated). Sequence parts encompassing T-DNA border regions and Tn5393 inverted repeats (IR) are shown below. The vertical arrows point to the junction between the right borders (RB) of T-DNA copies 1 and 2. Underlined is the 21 bp inverted repeat around the RB junction, originating from T-DNA inverted repeat insertion. The 81 bp terminal IR of transposon Tn5393 is marked with a box. 8 bp repeats originating from the second T-DNA copy upon insertion of Tn5393 are indicated by dashed lines. Stop codons of the tnpA gene within the 81 bp IR are shown in bold

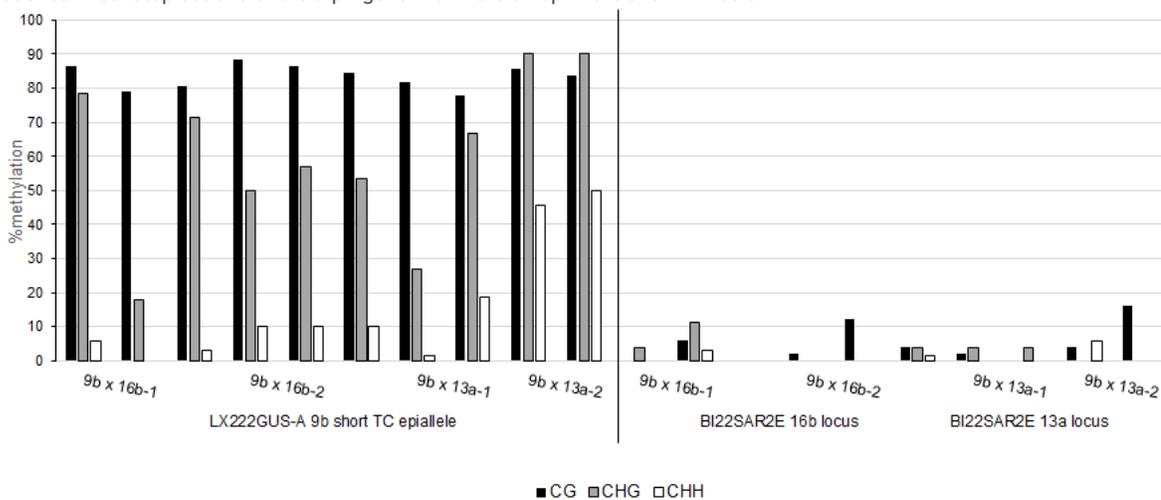


Figure 4

Differential DNA methylation analysis in the 3'gus-Tnos region of hybrids containing the LX222GUS-A 9b *sh* or *tTC* epiallele and one of the BI22SAR2E loci. Methylation percentages at CG, CHG and CHH sites were determined in 4 individual hybrid plants. Ten cloned PCR products were analyzed each for the LX222GUS-A 9b locus and for the transgene construct present in transgenic lines BI22SAR2E 13a and BI22SAR2E 16.