10. ATTACHMENTS

Appendix is not publicly available in the electronic version of the thesis, because it contains published works with copyright restrictions and results which are being published.

Appendix consists of the following files:

10.1. Publication 1

Tyč, D., Nocarová, E., Sikorová, L., Fischer, L., 2017. 5-Azacytidine mediated reactivation of silenced transgenes in potato (Solanum tuberosum) at the whole plant level. Plant Cell Reports, 36, pp. 1311–1322. DOI: 10.1007/s00299-017-2155-7

Publication is available on the website https://link.springer.com/article/10.1007/s00299-017-2155-7.

10.2. Supplementary material for Publication 1

All files are available on the website https://link.springer.com/article/10.1007/s00299-017-2155-7.

10.3. Publication 2

Manuscript (attached bellow) submitted to the BBA – Gene Regulatory Mechanisms, but the final decision has not been made to the date of the thesis submission.

10.4. Supplementary material for Publication 2

Supplement attached bellow.

10.5. Publication 3

Přibylová, A., Čermák, V., Tyč, D., Fischer, L., 2019. Detailed insight into the dynamics of the initial phases of de novo RNA-directed DNA methylation in plant cells. Epigenetics & Chromatin, 12, pp. 1–14. DOI: 10.1186/s13072-019-0299-0

Publication is available on the website https://epigeneticsandchromatin.biomedcentral.com/track/pdf/10.1186/s13072-019-0299-0.

10.6. Supplementary material and sRNA data sets of Publication 3

All files are available on the website https://epigeneticsandchromatin.biomedcentral.com/articles/10.1186/s13072-019-0299-0#additional-information.

The sRNA data sets used in this study are available in European Nucleotide Archive PRJEB32154, http://www.ebi.ac.uk/ena/data/view/PRJEB32154.

Unexpected variations in posttranscriptional gene silencing induced by differentially produced dsRNAs in tobacco cells

Vojtěch Čermák[1], Dimitrij Tyč[1], Adéla Přibylová[1], Lukáš Fischer[1*]
[1]Charles University, Faculty of Science, Department of Experimental Plant Biology, Viničná 5, Prague 2, 128 44, Czech Republic
[*]corresponding author: lukasf@natur.cuni.cz

Abstract

In plants, posttranscriptional gene silencing (PTGS) is induced by small RNAs (sRNAs) generated from various dsRNA precursors. To assess the impact of dsRNA origin, we compared downregulation of GFP expression triggered by inverted repeat (IR), antisense (AS) and unterminated sense (UT) transcripts transiently expressed from the estradiol-inducible promoter. The use of homogeneously responding tobacco BY-2 cell lines allowed monitoring the onset of silencing and its reversibility. In this system, IR induced the strongest and fastest silencing accompanied by dense DNA methylation. At low induction, silencing in individual cells was binary (either strong or missing), suggesting that a certain threshold sRNA level had to be exceeded. The AS variant specifically showed a deviated sRNA-strand ratio shifted in favor of antisense orientation. In AS lines and weakly induced IR lines, only the silencer DNA was methylated, but the same target GFP sequence was not, showing that DNA methylation accompanying PTGS was influenced both by the level and origin of sRNAs, and possibly also by the epigenetic state of the locus. UT silencing appeared to be the least effective and resembled classical sense PTGS. The best responding UT lines behaved relatively heterogeneously possibly due to complexly arranged T-DNA insertions. Unlike IR and AS variants that fully restored GFP expression upon removal of the inducer, only partial reactivation was observed in some UT lines. Our results pointed out several not yet described phenomena and differences between the long-known silencer variants that may direct further research and affect selection of proper silencer variants for specific applications.

Highlights

- dsRNA formation affects speed, reversibility and transitivity of silencing
- Silencing is binary in weakly induced cells, either strong or missing
- DNA methylation accompanying PTGS is influenced by sRNA level and origin
- Methylation induced by sRNAs can differ even in loci with identical DNA sequence
- Silencing by unterminated sense transgenes likely do not differ from classical S-PTGS

Keywords

DNA methylation, dsRNA, PTGS, RNAi, siRNA, tobacco BY-2 cell line

1 Introduction

RNA interference (RNAi) is an important mechanism involved in the regulation of gene expression in eukaryotic cells. It is often used in research as a tool to downregulate expression of a selected gene. The key players in RNAi are small RNAs (sRNAs), which can be formed via several different pathways in plants, making plant RNAi a very complex phenomenon [1]. Generally, a double stranded RNA (dsRNA) is needed to induce the production of sRNAs in plants. These dsRNAs are recognized by DICER-LIKE (DCL) proteins, which process them into sRNA duplexes. One strand of the sRNA duplex (guide strand) is then selected to associate with ARGONAUTE (AGO) proteins, while the other strand, so-called passenger strand, is degraded. AGO with a loaded guide strand recognizes various target RNA molecules by sequence complementarity. Four DCL genes and ten AGO genes are encoded in the *Arabidopsis thaliana* genome. The DCLs differ according to what type of dsRNA they prefer to dice and the length of the sRNAs they produce [2–5]. The AGO proteins bind sRNAs based on their length and the identity of their 5’ nt and select the guide strand by a thermodynamic mechanism in cooperation with DRB proteins [6–10]. AGO proteins can slice the RNA target or cooperate with other proteins to direct various processes like blocking translation or

directing DNA methylation (which involves a wide range of proteins e.g. SUO or DRM2, respectively) [11,12]. When the target is an mRNA molecule, then these processes lead to posttranscriptional gene silencing (PTGS). However, the same pool of sRNAs can also induce chromatin modifications (primarily DNA methylation), which can accompany the PTGS [13]. Such methylation does not significantly affect gene transcription, because it is usually limited to the transcribed region [14]. When this methylation spreads to the promoter region, then the PTGS can change to silencing at the transcriptional level (TGS; [15]).

In plants, DNA can be methylated at cytosines in both symmetrical (CG and CHG) and asymmetrical (CHH) contexts, which principally differ in the mechanism of their maintenance in newly synthesized strands after replication [16,17]. Hemimethylated CG sites can be easily recognized and maintained by METHYLTRANSFERASE 1 (MET1) in cooperation with VIM proteins [18]. However, additional signals from histones are needed for non-CG (CHG and CHH) methylation maintenance. DNA methylation is mutually linked to histone modifications through chromatin modifying enzymes, whose activity is either directly or indirectly controlled by the other type of the chromatin mark. Thus, CHG and CHH methylation is maintained in heterochromatin by CHROMOMETHYLASE 2 (CMT2) and CHROMOMETHYLASE 3 (CMT3) attracted by histone modification (H3K9me2) recognized directly by chromodomains of these enzymes [19,20]. Vice versa, SRA domains of H3K9 histone methyltransferases such as KRYPTONITE (KYP, SUVH4) and SUVH5/6 recognize DNA methylation [21–23]. The other signal for maintenance methylation is presence of hc-siRNAs that are involved in the canonical RNA-directed DNA methylation (RdDM) pathway. These hc-siRNAs are generated from products of RNA polymerase IV that transcribes chromatin with specifically modified histone H3 (H3K4me0, H3K9me2) which is recognized by SHH1 [24]. DNA methylation in RdDM is introduced by DOMAINS REARRANGED METHYLASE 1/2 (DRM1/2) when specific AGO protein carrying sRNA binds to complementary nascent scaffold transcripts generated by plant-specific RNA polymerase V (Pol V) specifically

attracted to methylated DNA by SUVH2/9 [12,25]. In addition to this maintenance methylation, RdDM also establishes *de novo* DNA methylation utilizing sRNAs of various origin including those arising during PTGS [13,26].

There are three basic ways how dsRNA can be formed. First, separate transcription of complementary (sense and antisense) RNA strands leads to their intermolecular pairing, generating dsRNA. This occurs naturally in the pathway generating natural-antisense siRNA (nat-siRNAs) [27]. The second way is based on the transcription of an inverted repeat; such a transcript creates a hairpin dsRNA through intermolecular pairing. Naturally, miRNAs and some other types of hairpin siRNAs are generated in this way [4,28]. The third way uses RNA-DEPENDENT RNA POLYMERASE (RDR) to convert certain ssRNAs into dsRNAs. This mechanism is naturally involved in the production of hc-siRNAs, ta-siRNAs and most of the secondary siRNAs [2,29,30]. It is also involved in the sense transgene silencing (S-PTGS) and can be connected with the phenomenon known as cosuppression [31] that is believed to result from high production of aberrant RNAs. These RNAs can saturate the RNA degradation pathway leading to their accumulation in the cell, which makes them available as a substrate for RDRs [32].

Many types of silencer constructs have been designed based on this knowledge and have been meticulously compared in an attempt to reach the highest level of silencing and the maximum percentage of silenced transformants (e.g. [13,33–35]). We hypothesized that the use of constructs based on different ways dsRNA forms would lead to the activation of different RNAi pathways and thus results in a different course of silencing, even if lines with similar decrease in expression are selected. The understanding of all the effects dsRNA formation has on the silencing outcomes is not only important for understanding the RNAi itself, but it is also important for the use of appropriate silencing constructs in applied research.

In our study, we chose the three most distinct ways dsRNA forms and prepared the following silencer constructs: antisense (AS), unterminated (UT; expected to produce aberrant RNAs; [32]) and

inverted repeat (IR); all of them based on (and targeting) the full-length coding sequence of *GREEN FLUORESCENT PROTEIN* (*GFP*) gene. We used an inducible promoter for these constructs to eliminate spontaneously silenced lines and to be able to observe the silencing from the very beginning to the steady state and during its termination (recovery of GFP expression). We supertransformed each of these constructs to the highly homogeneous tobacco BY-2 cell line [36] where they targeted *GFP* transcripts stably expressed from the identical locus in all the variants [37–39]. This allowed us to make direct comparisons of the effects of the three silencer variants. Although the model system based on cell lines and silencing of the *GFP* transgene was quite artificial, it allowed to control most of the variables so the observed differences could be attributed directly to the way dsRNA formed. With a similar system, we recently described the dynamics of the initial phases of *de novo* RdDM [39]. In this current work, we first compared dynamics of PTGS onset at the fluorescence level in populations of hundreds of independently transformed lines (calli). Thereafter, the silencing dynamics and recovery at mRNA, sRNA and DNA methylation levels and by flow cytometry at the single cell fluorescence level were analyzed in detail in selected lines with the highest silencing rate.

2 Results

2.1 Establishment of lines for inducible silencing and their comparisons at the population level

Three different silencer constructs were prepared to induce GFP silencing (Fig. 1). All were based on the full-length coding sequence of *GFP* gene and their transcription was controlled by β-estradiol inducible promoter (XVE system; [40]). The constructs were as follows: antisense *GFP* (AS), inverted-repeat *GFP* (IR) and unterminated *GFP* (UT) where the sense transcripts were expected to be aberrantly terminated and recognized by RDR (RNA-dependent RNA polymerase) to produce dsRNA. As a control in some experiments, we also used a sense *GFP* gene with a terminator (GF) that can under certain occasions undergo sense transgene posttranscriptional silencing (S-PTGS) connected with cosuppression of the target GFP. In cases where the *GFP* sequence was in the sense

orientation (UT and GF variant), it was without the start codon (unless otherwise stated) to prevent potential unwanted translation of the GFP from the silencer RNA.

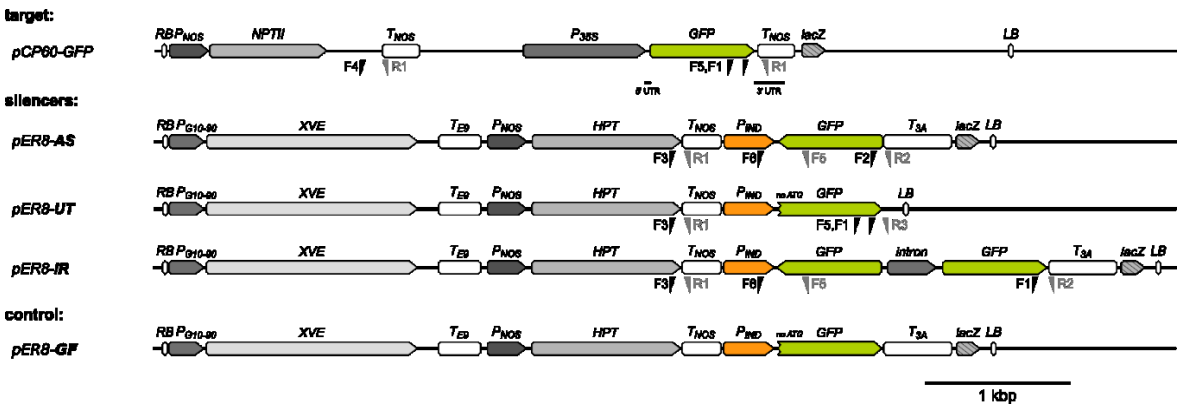


Fig 1. Scheme of the silencer and target *T-DNAs*.

Functional elements and positions of primers used for reverse transcription, qPCR and McrBC assay are shown at the appropriate sequences (primers itself are listed in Table S1). Target *T-DNA* of *pCP60-GFP* (https://www.addgene.org/122173) confers constitutive GFP expression. *pER8*: a set of *T-DNAs* used for the inducible expression: *pER8-AS*: *GFP* gene in antisense orientation; *pER8-UT*: sense *GFP* gene without terminator and without start codon; *pER8-IR*: sequence of *GFP* arranged as an inverted repeat; *pER8-GF*: sense *GFP without start codon and* with terminator as a control. *RB*: right border, *LB*: left border, P_{NOS}: nopaline synthase promoter, P_{35S}: *35S* promoter, $P_{G10\text{-}90}$: synthetic constitutive promoter; P_{IND}: inducible promoter activated by β-estradiol; T_{NOS}: nopaline synthase terminator; T_{E9}: *rbcS E9* terminator; T_{3A}: *rbcs S 3A* terminator; *XVE*: estrogen receptor and transcriptional activator; *NPTII*: kanamycin resistance gene; *HPT*: hygromycin resistance gene; *GFP*: green fluorescent protein coding sequence; intron: intron from Solanum tuberosum *PsbO* gene; *LacZ*: fragment of bacterial β–galactosidase gene. All *T-DNAs* are at the same scale as indicated with 1kbp scale bar.

We transformed these constructs into the BY-2 cell line with stable expression of GFP [37]. The transformants of IR showed high spontaneous silencing, which was an effect of the peculiar read-through transcription from adjacent genomic DNA (and possibly also upstream parts of the T-DNA), as shown in our previous study [38].

From the set of independently transformed lines, we chose those that showed sufficiently homogeneous and high GFP fluorescence (about one hundred for each variant) and planted one half of the callus on control medium (D) and the other half on induction medium with estradiol (E) and measured the fluorescence of those calli during 14 day period (Table 1; Fig. S1A,B; Table S2). From

these data, we analyzed the efficiency (percentage of lines able to induce silencing), strength (the maximum level of decreased expression) and speed of silencing. The IR was fastest in silencing induction and was able to induce the strongest silencing in a larger number of lines than any other variant. The ability of AS to induce silencing was somewhat lower and the UT was the least effective. The UT silencing did not significantly differ from the cosuppression caused by the control GF (Table 1), both variants contained only very few effectively silenced lines. Although the construct without terminator should produce aberrant RNAs, it was not always the case. Surprisingly, about 50% of lines, where an UT silencer construct with the start codon was transformed into the wild type BY-2 cells, were able to produce some GFP fluorescence (Fig. S1C). This indicates that in these lines the RNA was at least partially polyadenylated possibly due to T-DNA insertion in front of an endogenous terminator or due to a specific T-DNA arrangement. Otherwise, the mRNA could not be exported from the nucleus and translated [41].

variant	% of silenced lines			# of analyzed lines
	0 days	3 days	10 days	
AS	0.0	0.9	11.6	112
UT	0.0	2.0	9.0	100
GF	1.0	1.0	9.1	99
IR	1.6	30.6	45.2	62

Table 1. Efficiency of GFP silencing in BY-2 calli.

Percentage of lines with inducible silencer constructs or with the control sense *GFP* (GF) that were able to induce strong silencing of GFP expression (reduction of GFP fluorescence below 50% of mock-induced control) at different time points after the transfer to β–estradiol containing medium (lower number of available IR calli was caused by the high frequency of spontaneous silencing). Statistics for this data are in Table S2.

2.2 Detailed analyses of the time course of silencing

For further experiments, we selected three lines of each variant (named IR1-3, AS1-3 and UT1-3) with highly homogeneous *GFP* fluorescence (determined by flow cytometry at the single cell level) and with the highest silencing rate, as such lines are preferentially used in reverse genetics. Comparing lines with the potential to fully induce silencing also allowed us to look for specific differences between the silencer variants that were not directly dependent on the differences in silencing strength between the variants.

The estradiol inducible system enables to adjust the strength of silencer transcription by changing the concentration of estradiol in the medium. We first tested the GFP silencing in calli of the selected lines on 2μM concentration (commonly used in the XVE system) and a series of dilutions (Fig. 2). The IR showed the fastest and strongest response, even on 16nM concentration of estradiol. For the IR variant, the effect of the estradiol concentration was most prominent at the final state of silencing, rather than at the speed of silencing. Interestingly, the UT silencer was also able to induce silencing even on 16nM estradiol when given enough time, and the differences in the strength of silencing were the least dependent on the estradiol concentration in this variant. In contrast, the AS variant showed a strong dependence on the estradiol concentration, with 16nM concentration being too low to induce any significant silencing (Fig. 2).

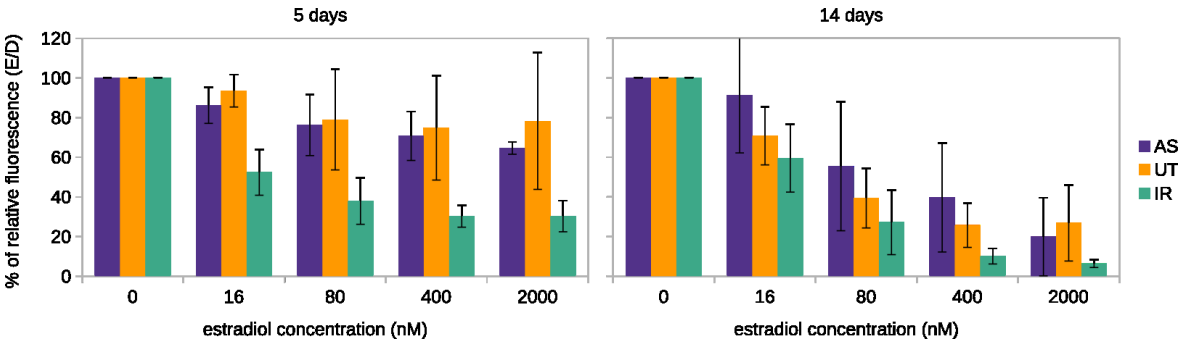


Fig 2. Effects of estradiol concentration on the induction of silencing.

The ratio of relative GFP fluorescence between calli cultured on medium with indicated estradiol concentrations and without estradiol (E/D) for 5 and 14 days. Data were normalized for day 0 to be 100%. For each silencer (AS, UT, IR) three lines were tested and for each line three calli were measured. The error bars represent standard deviation.

The previous experiments suggested that the standard estradiol concentration (2μM) was saturating for most lines, so it was used in the subsequent experiments. The three selected lines for each variant were transferred to suspension cultures and exposed for 14 days to estradiol and then the estradiol was washed out and the cells were kept for additional 21 days on the standard medium. Control cells for each line were handled in the same way, but without exposure to estradiol.

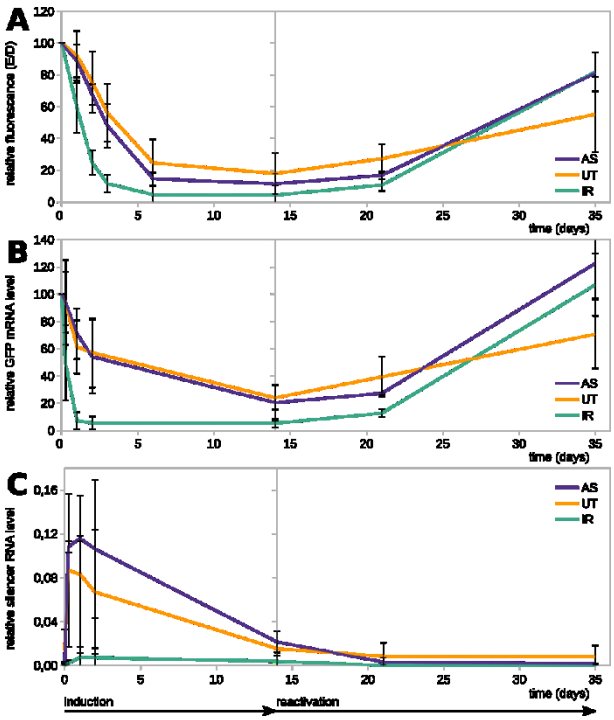


Fig 3. Dynamics of GFP silencing and recovery in suspension cultures.

Three best responding lines for each silencer variant were treated with β–estradiol for 14 days, then β–estradiol was washed out and the cells were monitored for an additional 21 days. (**A**) Time course of GFP fluorescence as measured by flow cytometer. The relative fluorescence is shown as the ratio of fluorescence between induced and mock-induced cells (E/D). Data were normalized for time 0 to be 100%. Data for AS2 on the 35th day are not available. (**B**) Time course of *GFP* transcript levels as measured by RT-qPCR. The relative *GFP* transcript levels were normalized to the internal standard *EF1α* and related to time 0 that was set to 100%. (**C**) Time course of silencer transcript levels as measured by RT-qPCR. The relative silencer transcript levels were normalized to *EF1α*. Error bars in all plots represent standard deviation. Data for each line individually are shown in Fig. S2.

Based on the fluorescence measurements using a flow cytometer (Fig. 3A; Fig. S2A,D) silencing in the IR lines was already obvious on the first day and by the 14th day, the fluorescence decreased to 1-11% of the original fluorescence level, reacting almost completely homogeneously. After removal of estradiol, the fluorescence mostly returned to the original level. The AS lines showed the first signs of a reaction later, on the second day, and by the 14th day decreased to 6-20% of the original level. This reaction was mostly homogeneous, and fluorescence returned almost to the original levels after removal of the inducer. The reaction of UT lines was more heterogeneous compared to other variants with the UT1 line resembling IR lines (Fig. S2A). On the 14th day, the fluorescence decreased to 6-30% of the original level. After removal of the inducer, many cells of UT3 and UT2 line were able to maintain the GFP silencing (Fig. S2D).

Changes in RNA transcript levels were faster than changes in fluorescence (Fig. 3B; Fig. S2B), probably due to the long lifetime of the GFP protein [39]. At the transcript level, all IR lines reacted within the first six hours after the exposure to estradiol ($p < 0.05$, paired t-test), reaching 20-74% of the original transcript level (Fig. S2B). The silencing reached its maximum as early as in 1-2 days (2-11%). Only one AS line and one UT reacted so rapidly within six hours. The responses of the other lines were delayed; the first decreases were detected after one day. Silencing in AS and UT lines progressively grew during the whole treatment and reached 12-36% and 17-34% after 14 days for the AS and UT lines, respectively.

Surprisingly, transcript levels of the silencer were highest between 6 and 24 h for AS and UT but peaked between 24 and 48 h for IR (Fig. 3C; Fig. S2C). Moreover, the IR silencer transcript levels were about 10-fold lower than those for AS and UT (apart from the UT1 line).

2.3 IR silencing under weak induction.

The IR variant showed stronger response than the other variants, so we analyzed its course of silencing in more detail on estradiol concentrations near the expected threshold for silencing induction: 0, 4, 16 and 64nM. Flow cytometry data of the average fluorescence and RNA expression data showed that two IR lines were sensitive enough to induce silencing, even on 4nM estradiol, whereas the IR3 line began to react on 16nM estradiol (Fig. S3A, B). This indicates that these concentrations were really at the lowest edge for silencing induction. At these concentrations, the silencing in individual cells (seen on flow cytometry data in Fig. 4, Fig. S3C) was not homogeneous (continuous in the strength of response) but was present in two distinct states – active and silenced.

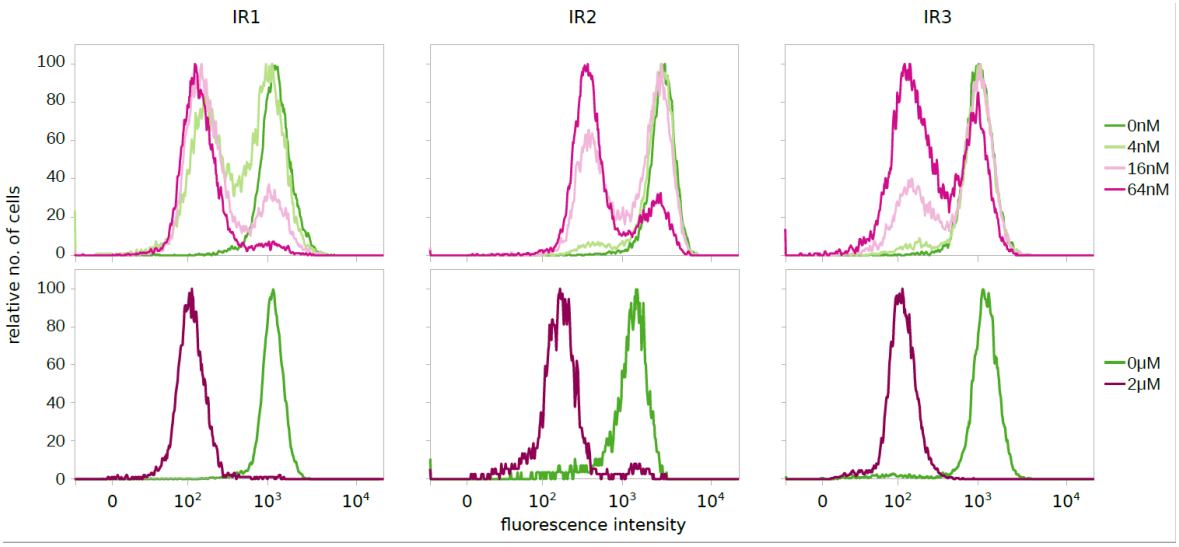


Fig 4. Effects of estradiol concentration on the induction of silencing after 3 days.

Histograms of cell GFP fluorescence levels (flow cytometer data normalized to mode) showing binary silencing at different estradiol concentrations in three different IR lines. The second line of graphs shows data for fully induced lines (2µM estradiol) from the experiment documented in Fig. 3.

2.4 Molecular characterization of the best responding UT lines

The behavior of the UT1 line compared to the UT2 and UT3 was abnormal and resembled silencing induced by IR (stronger and faster onset, almost complete reversion of silencing after removal of estradiol, and very low silencer transcript levels). It led us to examine the arrangement and copy number of *T-DNAs* in the three UT lines, because UT *T-DNAs*, when inserted as tail-to-tail inverted repeats (LB to LB), might produce a hairpin RNA. The number and size of hybridizing restriction fragments on Southern blot (a higher number of fragments detected by *GFP* probe after *Kpn*I digestion; Fig. S4) indeed suggest the presence of such an inverted repeat in the UT1 line, though we were unable to confirm its presence by PCR (data not shown). But it is known that PCR over inverted repeats is problematic [42]. Southern blot analysis did not indicate the presence of inverted repeats in UT2 and UT3 lines, but it suggested the presence of direct-tandem repeat in the UT3 line and indeed we were able to confirm it by PCR (Fig. S4E) and sequence the *T-DNA* junction (Fig. S4F).

The larger number of detected *T-DNA* insertions raised a question of the possibility of read-through transcripts at the *T-DNAs*, either from other *T-DNAs* in case where they would be in the repeat arrangement or from genomic DNA, resembling results from our previous study [38]. Indeed, we were able to detect low levels of both sense and antisense RNAs over different regions of the silencer in all three UT lines (Fig. S4G, H). Some of these RNAs were present already before induction, while others appeared in the response to estradiol induction. This behavior could be explained by the T-DNA arrangement as a result of readthrough transcripts from neighboring T-DNAs (in UT1 and UT3 lines) or by the activity of RDRs, as antisense RNA levels from genic regions behaved differently compared to intergenic regions. From highly expressed genic regions targeted by sRNAs, RDRs were shown to synthesize detectable amounts of complementary RNAs [30].

2.5 Analyses of small RNAs

The key determinant of PTGS are sRNAs, so we sequenced their populations in lines with a typical course of silencing for IR and AS variants (IR1 and AS3). Due to heterogeneity in UT lines, we could not select a really typical line, so we decided to analyze the line UT3, which was interesting in that it remained strongly silenced even after estradiol removal. Samples for this analysis were collected during the experiment analyzing the time course of silencing. Sequenced sRNAs were mapped on the target and the silencer *T-DNAs*. Some low levels of sRNAs aligning to the target sequence were detected even before induction (especially in the case of IR; Fig. 5A) which might have originated from a weak spontaneous transcription of the silencer [38], although it seemed that they had no effect on silencing. Estradiol induction led to the accumulation of target-specific sRNAs in all variants in quantities reaching 10 to 100 times the original level. The target T-DNA-specific sRNAs accounted for 0.5 to 1.9% of all sRNAs sequenced from the cells (Fig. 5A; Fig. S5A, B; Table S3A). After removal of estradiol, the high levels of sRNAs detected in IR declined, even beyond the initial levels. In contrast, target-specific sRNAs did not decrease in the UT3 line, but even increased further (Fig. 5A), which could explain the maintenance of silencing observed in UT lines after estradiol removal (Fig. 3A, B; Fig. S2A, B; data for AS at day 35 are not available).

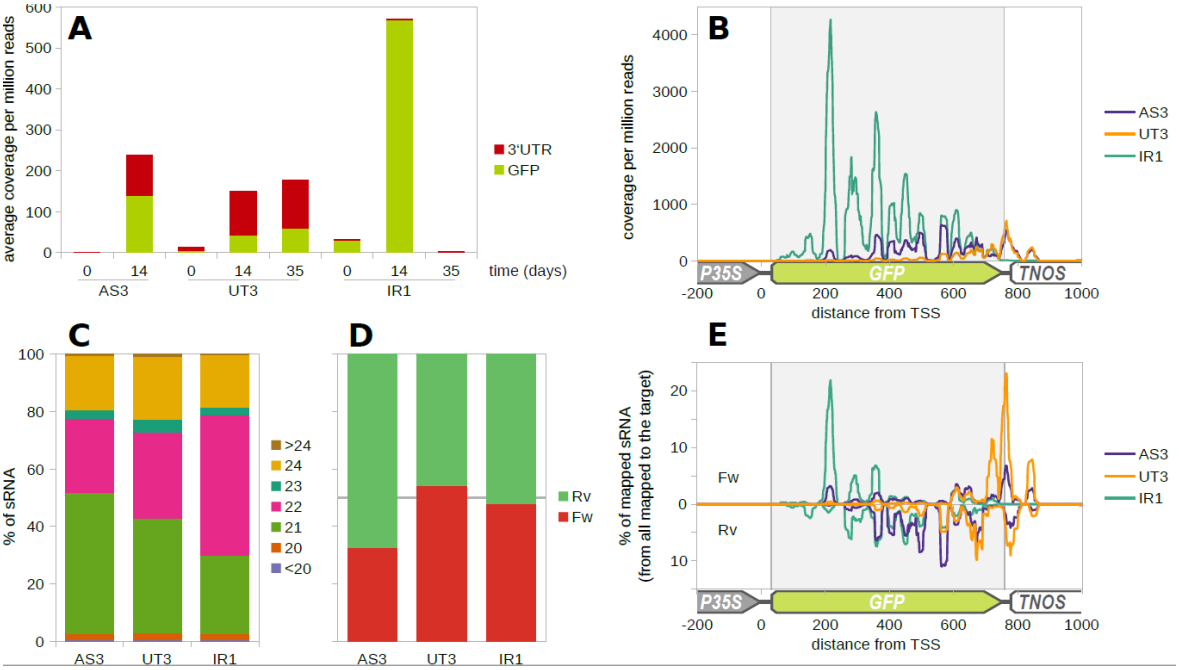


Fig 5. Characterization of sRNAs matching with the target *T-DNA*

(**A**) Average coverage of the target *GFP* and its 3' untranslated region (*3'UTR*) with sRNAs. (**B**) Distribution of sRNAs of all size categories along the target *GFP* and its 3' untranslated region at day 14 (more detailed analyses are presented in Fig. S5G.). (**C**) Length distribution of sRNAs matching to the target *T-DNA* at day 14. (**D**) Strand distribution of sRNAs matching to the target *T-DNA* at day 14 (Fw = forward, Rv = reverse). (**E**) Distribution of sRNAs of all size categories along the target *GFP* and its 3' untranslated region at day 14 based on their strand orientation (Fw are above X axis, Rv are below X axis). Each line is on a different scale for easier comparison of the position of individual peaks (more detailed analyses are presented in Fig. S5E).

Target T-DNA-specific sRNAs: At the target *T-DNA*, almost all sRNAs matched with the *GFP* transcription unit (more than 91.4%; Fig. S5A). In the case of IR, the sRNAs almost exclusively matched with only the *GFP* coding sequence, whereas in the case of AS and UT, there was strong 3' transitivity to the terminator *3'UTR* region (Fig. 5A, B). No obvious 5' transitivity at the target *T-DNA* was seen in any of the samples (Fig. S5C).

The silencer variants also differed in the representation of sRNAs of different lengths, more specifically, in the ratios of 21/22 nt sRNAs (Fig. 5C; Fig. S5E). The AS and UT had a dominant fraction of 21 nt sRNAs (49 and 40% respectively), while 22 nt sRNAs dominated in the IR (49%). The levels of 24 nt sRNAs were similar among the variants (around 20%) and the other sRNA sizes

made up only a small fraction (less than 8% altogether).

The ratio of sRNAs aligning to the forward and reverse strand of the target *GFP* gene was around 50% for IR and UT variants, but the AS variant deviated, with levels of the reverse strand sRNAs twice as high compared to forward strand sRNAs (Fig. 5D; 5: Fig. S5E). The variants did not show any larger differences in the representation of the 5' nucleotide of the target specific sRNAs (Fig. S5F).

sRNAs mapping to the target *GFP* transcription unit were not distributed equally (Fig. 5B, E; Fig. S5G), but they formed distinct peaks on either the reverse or forward strand. These peaks had highly similar positions between different sRNA size classes. Positions of peaks in the same region were also similar between different variants, but the variants differed in the heights of individual peaks and their representation along the sequence. In the IR variant, most of the sRNAs came from the 5' part of the *GFP* gene; in the AS, the distribution of peaks was more or less homogeneous along the whole transcription unit (including *3'UTR*); and in the UT, most of the sRNAs came from the 3' end – actually 41% of sRNAs in UT at the GFP transcription unit come from the *3'UTR*, which was not primarily targeted by the silencer. In the AS line, high production of secondary sRNAs can also be inferred thanks to a point mutation in the silencer GFP sequence introduced during cloning (A to T at 404. nucleotide from the start of GFP cds, see Fig. S5H for the resulting AS silencer sequence). This single point mutation was expected to have no influence on the silencing, but it could be used to partially distinguish between primary and secondary sRNAs; sRNAs where the antisense strand lacks the mutation or where the sense strand has the mutation cannot be primary sRNAs. Considering this criterion, at least 60% of sRNAs overlapping this nucleotide position had to be secondary (Table S3B).

sRNAs specific to the silencer T-DNAs: In the case of the silencers, there were more regions outside the primary sRNA source sequence that showed accumulation of sRNAs after the estradiol induction

(Fig. S5B). As with the target *GFP*, the AS and UT silencers showed 3' transitivity (although weaker), but no obvious 5' transitivity (Fig. S5C). Interestingly in the UT silencer, which has no terminator, the levels of sRNAs sharply decreased at the end of the *GFP* sequence and showed smaller accumulation of sRNAs at the region downstream of the *GFP* coding sequence than in the terminated AS silencer (Fig. S5B, C).

Beside sRNAs matching the silencer transcripts, there were also sRNAs from the *HPT* (hygromycin phosphotransferase) transcripts in the AS and UT variants (Fig. S5B, C, D). The *HPT* gene shares *3'UTR* with the target *GFP* gene. So, the 3' transitive sRNAs from the target *T-DNA* could also target the *HPT* gene and cause 5' transitivity into the *HPT* coding sequence (this was the only clear case of 5' transitivity in our experimental system). This was also supported by the decrease in *HPT* mRNA levels in AS and UT lines after the induction of silencing (Fig. S5I). The *NPTII* (kanamycin phosphotransferase) gene in the target *T-DNA* also shares the same *3'UTR* as the target *GFP* gene and the *HPT* gene, nevertheless, there were no secondary sRNAs spreading into the *NPTII* gene (Fig. S5A, C). However, like with the *HPT* gene, mRNA levels of the *NPTII* gene also decreased when the lines were exposed to estradiol (Fig. S5I). In the UT3 line, there were also high levels of sRNAs matching to the 3' part of the *XVE* gene encoding the estradiol receptor (Fig. S5B). Most of these sRNAs appeared only after induction and they behaved the same as the *GFP* sRNAs. These sRNAs were probably related to the presence of the direct tandem repeat in this line, as described above (Fig. S4).

In the IR variant, the amounts of sRNAs that matched regions outside the *GFP* sRNA source sequence were extremely low (Fig. S5C). One of the regions with low levels of sRNAs was the linker between the inverted *GFP* sequences. When analyzing the sRNA data, we uncovered that the line IR1 has a rearrangement in the linker between *GFP* sequences, which is much larger than it should be (Fig. S5J). Sequencing of the linker showed that the sequence between the exact end of the first *GFP* in the repeat to the donor splice site was replaced with 839 bp from the central part of the *XVE* gene and 1514

bp from the backbone of pER8 plasmid, starting with the end of the left border region (Fig. S5J). Though unintended, the fact that the linker contained part of the sequence that was not transcribed from any other locus and part of a sequence that was transcribed within the *XVE* transcription unit, gave us an opportunity to analyze the sRNAs from this region in more detail (Fig. S5K, L). The sRNAs from the region homologous to the *XVE* reached higher levels (3-fold) than those in the surrounding regions. They also differed in the representation of sRNAs of different lengths (32% of 22 nt from region matching *XVE*, compared to 45% of those from the surroundings; Fig. S5L). Even though the levels of sRNAs from the linker reached only low levels, they were able to induce weak 3' transitivity on the *XVE* transcript (Fig. S5C).

2.6 DNA methylation accompanying the PTGS

Presence of sRNAs is known to induce DNA methylation. To see whether our silencers differed in this aspect we performed bisulfite sequencing of two regions encompassing approximately 600 bp from the *35S* promoter and 600bp from the *GFP* coding sequence (Fig. 6A, B; Fig. S6A, B) at three key time points in the experiment (0, 14 and 35 days). There was no DNA methylation during the whole experiment in the AS variant. In contrast, the IR was able to induce high levels of DNA methylation in all sequence contexts along the whole analyzed target *GFP*. Specifically, the methylation in the CHG context reached 96% and when looking at the CWG context alone the methylation was 100% (there were 14 CWG cytosines in that region). There was also a weak spill-over of methylated cytosines outside of the target *GFP* at the 5' end, that reached about 40 nt upstream (we have no data from the 3' end of the target *GFP*). After removal of estradiol, the methylation in the CG context remained at the same level, but the methylation in CHG and CHH context almost fully disappeared (Fig. 6A; Fig. S6B).

In the UT3 line we chose for detailed analyses, DNA was methylated in the CG context already at time 0, which may reflect some transient silencing that might have occurred in this particular line before the experiment began. After induction, there was a weak increase in CHG and

CHH methylation at the 3' end of the *GFP* sequence, but a decrease in CG methylation in the middle of the target *GFP* (the changes in both CG and CHH were statistically significant, $p < 0.05$).

Though our previous study showed that preexisting CG methylation did not influence *de novo* methylation [39], we included one additional line (UT2) to our analysis. This line was free of the initial CG methylation, and the estradiol induced methylation in all three sequence contexts. This methylation was limited to the 3'end and remained almost unchanged even after estradiol removal (Fig. S6C).

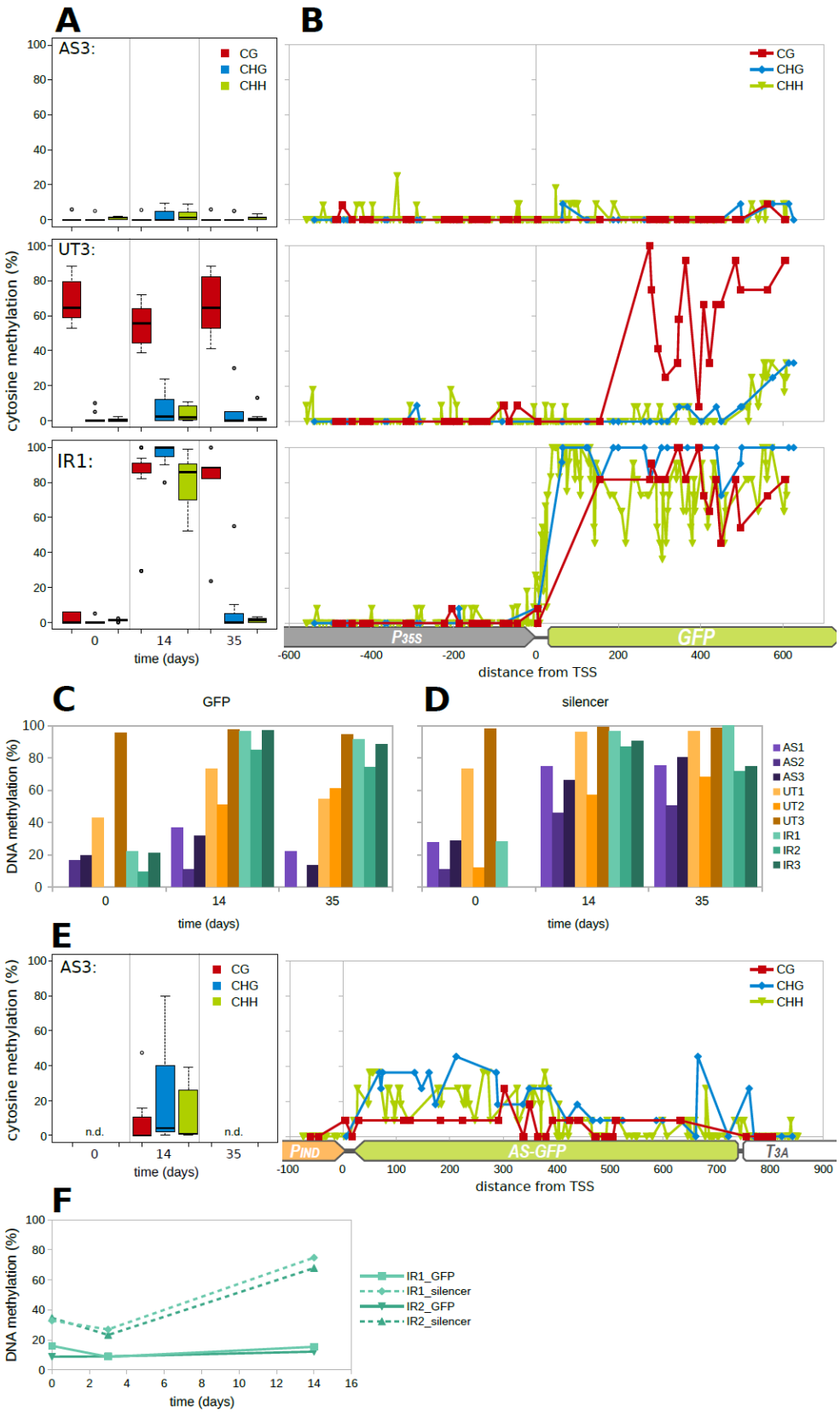

A
AS3:
UT3:
IR1:
CG
CHG
CHH
cytosine methylation (%)
time (days)
B
distance from TSS
P35S
GFP
C
GFP
DNA methylation (%)
time (days)
D
silencer
AS1
AS2
AS3
UT1
UT2
UT3
IR1
IR2
IR3
E
n.d.
PIND
AS-GFP
T3A
F
IR1_GFP
IR1_silencer
IR2_GFP
IR2_silencer

Fig 6. DNA methylation of the GFP target and silencer during PTGS

DNA methylation levels were assessed in suspension cultures treated with β–estradiol for 14 days and recovered for an additional 21 days. (**A**) Boxplots showing average DNA methylation levels in CG, CHG and CHH contexts in the analyzed region of the target *GFP* gene. Analysis was done by bisulfite modification in one line representing each silencer variant (lines AS3, UT3 and IR1), 8 to 12 DNA clones were sequenced for each data point. (**B**) Distribution of methylated cytosines in the three sequence contexts along the target region and in the 5' adjacent region after 14-day β–estradiol treatment. (**C**) Methylation level at the target *GFP* gene and (**D**) at the *GFP* region of the silencer *T-DNA* assessed by McrBC assay (primers for qPCR are listed in Table S1). (**E**) Boxplots showing levels of DNA methylation in CG, CHG and CHH contexts for the silencer *GFP* region in AS3 line at 14th day of induction, and distribution of methylated cytosines along the silencer *GFP* region (including 3' and 5' adjacent regions) for the same data (data from 11 sequenced DNA clones). (**F**) Target GFP and silencer methylation over time as assessed by McrBC assay for lines IR1 and IR2 at 16nM concentration of estradiol.

To verify the results of the bisulfite sequencing in all three lines of each variant, we used McrBC assay that is based on qPCR quantification of DNA digested with an enzyme that specifically cuts methylated DNA. The results showed that methylation in all the AS lines was very weak or missing. In contrast, all three IR lines showed strong induction of methylation that remained high even after estradiol removal, likely in the form of CG methylation (Fig. 6C). The UT lines differed among themselves. As expected, the UT3 line with preexisting CG methylation showed consistently high methylation, while the UT2 line showed estradiol-induced methylation and the UT1 line behaved somehow in between.

We also used the McrBC assay to analyze the methylation at the silencer locus. For the UT and IR variants, the behavior was highly similar to the situation at the target locus, but the methylation differed in the case of the AS variant (Fig. 6D). Unlike the target region, the AS expression was able to induce some DNA methylation at the silencer locus. To verify this result, we did bisulfite sequencing of the AS silencer on the 14th day after induction from the same sample that we used for the target *GFP* analysis (Fig. 6E). Indeed, we did detect mid-levels of DNA methylation, especially in the CHG context, but also weak methylation in the CHH and CG context along the whole antisense *GFP* sequence. Using the McrBC assay we could also see that detectable changes in methylation in the IR1 line appeared after two days at the target *GFP*, but as early as after one day at the silencer

locus (Fig. S6D). Moreover, it turned out that the weakly induced IR was not able to induce DNA methylation at the target locus, although it was still able to induce weak methylation at the silencer locus (Fig. 6F).

3 Discussion

In this study, we closely analyzed the effects of different silencers (sources of sRNAs) on PTGS, its strength, efficiency, timing, dynamics, reversibility, and accompanying processes like DNA methylation, secondary sRNA production and transitivity. All silencers were based on identical sequence, full-length *GFP*, and the silencing target was constitutively and stably expressed *35S::GFP* in the identical locus, allowing direct comparisons of the silencing features. Our results showed consistent behavior in all selected IR and AS lines indicating that their character was indeed connected with different sRNA sources. In case of UT silencer, the best responding lines were more heterogeneous possibly due to involvement of other RNAi pathways that might be related to complex T-DNA insertions in these lines.

3.1 Silencing strength, efficiency and dynamics differed among silencers

Our data confirmed previous observations that IR is the most effective inducer of silencing (Table 1; Fig. S1) most likely thanks to the highest efficacy of dsRNA formation leading to very high levels of sRNAs. IR silencer caused the strongest reduction of target gene expression (as low as 2% of the original transcript level; Fig. S2A, B; [13,33–35]). The efficiency that we observed for IR was lower than usually obtained, which likely resulted from the fact that about 70% of IR lines had silenced GFP expression spontaneously (before induction with estradiol) due to read-through transcription from the adjacent genomic region into the *T-DNAs* [38]. These lines, which were not included in our analysis, would have been probably classified as successfully silenced when assessing the efficiency with a constitutively expressed silencer. Silencing induced by UT [32,43,44] was less frequent in our

experimental system and did not significantly differ from classical cosuppression (induced by terminated sense *GFP* transcripts; Table 1; Fig. S1) challenging the impact of the missing terminator for silencing induction.

With data from the inducible system, we could extend the comparisons beyond just silencing strength and efficiency. The silencing with IR was faster when compared to UT and AS (with the first reaction within 6 h; Fig. 3; Fig. S2) similar to the 3 to 24 h that was previously observed using inducible silencing with IR in *Arabidopsis* and tobacco [45–47].

When testing the impact of the strength of silencing induction, we found that at the cell level the silencing was not continuous in the IR lines, but the cells were either silenced or not (Fig. 4). This might hint to some mechanisms that perhaps serve to prevent sporadically produced sRNAs from inducing unwanted silencing. Only when sRNA level reaches a certain threshold then they will have the potential to induce silencing [48].

In the UT lines, one would expect that silencing is strongly dependent on the intensity of induction with some threshold level because UT silencing is believed to be dependent on production of aberrant unterminated transcripts. Such transcripts could enter the RNAi pathway only after they saturate the RNA degradation pathways [32], so the UT construct should perform much worse or completely fail at low induction levels. Surprisingly, it performed much better than AS and, given enough time, it almost matched the IR (Fig. 2). This low dependence of UT silencing on the inducer level and the same performance of UT and GF lines in the population callus screen suggest that the mechanism of silencing in the UT lines might be different from that we expected. Recently, Parent et al., 2015 [49] documented that S-PTGS in *Arabidopsis* L1 line (line with spontaneous silencing of GUS transgene) was connected with the formation of aberrant read-through antisense RNAs from the oppositely oriented *NPTII* transgene. Though it might be considered as a specific situation, a similar mechanism seems to be important in the regulation of a large number of *Arabidopsis* genes [49]. Other studies also linked cosuppression to the read-through transcription [50] and some of the

important works on S-PTGS were done on plants with convergently oriented transgenes including the work of Luo and Chen, 2007 [32]. Our results show that in lines carrying the UT construct, aberrant untranslatable RNAs arose more frequently (by at least 30%) compared to the control terminated GF lines (Fig. S1C), but the frequency of silencing did not differ between the UT and GF variants (9% in both). It suggests that aberrant unterminated RNAs in our UT lines were probably not sufficient for the induction of efficient silencing. More likely, the transcripts generated from the inducible promoter required other features for silencing induction, perhaps similar to the situation described in the *Arabidopsis* L1 line, where silencing was connected with the antisense read-through transcription ([49]; Fig. S4G). This read-through transcription might result from the complex arrangement of *T-DNA* insertions or from the genomic context as discussed previously [38,49]. As such, it might be much more difficult or even impossible to induce such silencing without more complex T-DNA structure.

3.2 Silencing with UT could not be fully recovered

IR silencing in our experiments showed practically full transcriptional reactivation after the removal of the inducer consistently with previous studies [45–47]. The same was observed for the AS variant, but in the UT2 and UT3 lines GFP expression could not be fully recovered after estradiol removal (Fig. 3; Fig. S2). The silencing maintenance was unlikely caused by promoter leakiness because in all analyzed lines the silencer transcript levels at day 35 were lower than the background levels detected at day 0 (Fig. S2C). We also did not observe any DNA methylation spreading into the *35S* promoter (Fig. 6B), so it was unlikely that the persisting silencing was caused by TGS. PTGS maintenance without the inducer was already observed after viral infection [51] or transient agrobacterium transformation [52]. Also, classical S-PTGS that has features in common with UT silencing can be maintained when spontaneously initiated [53]. However, specific features of this silencing were, to our knowledge, never described in detail. Here we show that the maintenance of silencing in the UT3 line was accompanied by the continual presence of sRNAs that even slightly

increased in abundance after the inducer removal (Fig. 5A). Stable sRNA levels after inducer removal can also be inferred for the UT2 line from persisting CHH methylation. Interestingly, there were no evident differences between the profile and composition of sRNA for the UT and AS lines that could easily explain this difference in their behavior (Fig. 5B, E; Fig. S5B, C, E). The *3'UTR* sRNAs for the AS and UT lines were virtually identical and their ability to induce production of transitive sRNAs from the HPT gene was the same. There were some differences between the sRNAs matching to the *GFP* region, however, the UT sRNAs in this region seemed to be mostly a subgroup of sRNAs present in the AS line. As such, the reason for the maintenance of the UT silencing likely lies somewhere else, like in the source or in the site of sRNA production. Nevertheless, the inducible UT construct could provide a tool for studying the mechanisms involved in the maintenance of PTGS.

3.3 UT and AS silencing were accompanied with strong sRNA transitivity

Sequencing of sRNAs identified large amounts of secondary sRNAs in the AS and UT line. In the UT line, the transitive sRNAs against *3'UTR* made up 41% of the sRNAs targeting the *GFP* transcript (Fig. 5A, B). Moreover, after the removal of the inducer, there was no decrease in their amount and no change in their profile (Fig. 5A; Fig. S5E), suggesting that the primary sRNAs made up only a very small fraction of the sRNAs targeting the *GFP* transcript. Similar amounts of transitive sRNAs were also present in the AS line (Fig. 5A, B). Also based on a single point mutation in the AS silencer, secondary sRNAs made up at least 60% of sRNAs overlapping that position (Table S3B). In both the AS and UT line, the *3'UTR* secondary sRNAs were also able to induce production of secondary 5' transitive sRNAs in trans from the *HPT* gene, but interestingly not from the *NPTII* gene with an identical terminator (Fig. S5C). This could have resulted from different sequences present upstream of the T_{NOS} part of the *3'UTR* sequence, or it could have resulted from different distances between the sRNA recognition site and the end of the coding sequence between both genes (29nt for *HPT* and 387nt for *NPTII*; Fig. 1). It has been previously shown that translation and production of secondary sRNAs are closely connected [54,55]. The *HPT* gene is also the only clear example of 5' transitivity

in our system. This might be because in our system sRNAs mostly targeted coding sequences, where the 5' degradation products could be efficiently degraded by the nonstop decay (NSD) pathway [56], whereas in the case of *HPT* gene, sRNAs targeted the untranslated region.

3.4 IR silencing was not accompanied with transitivity despite high level of 22 nt sRNAs

While in the AS and UT lines there was strong 3' transitivity on the *GFP* transcript and the majority of sRNAs seemed to be secondary (Fig. 5A, B), in the IR line, there was no evidence of the production of significant amounts of transitive sRNAs. Only rare sRNAs from the linker between the arms of the inverted repeat were able to induce low levels of transitive sRNAs on the *XVE* transcript (Fig. S5K). Missing 3' transitivity is in contradiction with what has been shown previously that inverted repeats were able to induce strong 3' transitivity, especially when targeting transgenes [57–60]. Also, unlike the situation in our system, most inverted repeat sRNAs are dominated by 21 nt fraction of sRNAs, because DCL4 is the lead dicer in processing long dsRNA [3,60–62]. This is a characteristic also common in most endogenous IRs in *Arabidopsis*, except for IR71, which is dominated by 22 nt sRNA [63]. Dalakouras et al. observed silencing of GFP induced by an inverted repeat, where the majority of sRNAs were 22 nt long and there were only low levels of transitive sRNA [64], resembling our case (Fig. 5B, C). They speculated that this was caused by the high expression of DCL2 in *N. benthamiana*. The same might also apply to the *Nicotiana* BY-2 cells, but not generally, because when we expressed *CaMV 35S* promoter hairpin in BY-2 cells, the sRNA size classes were similar to other IRs, with the dominant fraction being 21 nt long [39].

3.5 AS silencing showed strong sRNA strand asymmetry

Data from sRNA sequencing showed an extremely shifted ratio of forward and reverse strand sRNAs in the AS line (Fig. 5D). The preference for the sRNAs to originate from the forward or the reverse strand is locally affected by the sequence composition, as can be seen from the distribution of peaks of sRNA coverage along the sequence (Fig. 5E). On a larger scale, the strand ratio is almost even for

the whole *GFP* transcript in the UT and IR lines. The shift in the ratio favoring antisense strands observed in AS line was mostly caused by decreased sRNA production from two regions with a strong bias for forward strand sRNAs at the 5' and 3' end of the *GFP* coding sequence compared to the IR and UT lines, respectively (Fig. 5E; Fig. S5G). The stronger production of sRNAs at the 5' end in the IR line was probably caused by the way the hairpin RNA is processed. It has been previously shown that more sRNAs are produced from regions closer to the loop [59], which is the 5' end of the *GFP* sequence in our case. However, at the 3' end the AS line seemed to be able to produce sRNAs well enough, but the production was lower only in one specific location, with the highest bias towards the forward strand, affecting all types of sRNAs (Fig. S5G). To be able to explain what caused this difference, it would be necessary to know which sRNAs are primary and which are secondary. Though the mechanism responsible for the strand bias remains unclear, it might be a general feature of processing dsRNAs that originate from intermolecular pairing, as it is common among nat-siRNAs to have a highly uneven sRNA strand ratio [65].

3.6 Loci with the same sequence can show different sensitivity to DNA methylation

The AS variant also showed unexpected differences in its ability to methylate corresponding DNA sequences. All the AS lines could clearly methylate the sequence of the silencer, but they were unable to induce significant changes in methylation at the target *GFP* (Fig. 6C, D). To explain this difference between methylation of the target and the silencer locus in AS lines, one might look for a connection with the strand asymmetry in sRNA production. But it is unlikely the cause since weakly induced IR lines behaved in the same way (Fig. 6F). So, we prefer an alternative explanation. The GFP from the target locus has been stably expressed for several years [37], whereas the expression from the silencer locus has been switched on right at the start of the experiment. It is reasonable to expect that these loci differ in their epigenetic state. The locus with more chromatin activation marks could be more resistant to deposition of repressive marks, as was reported in an animal system [66]. Alternatively, the stronger ongoing transcription could more effectively erase newly deposited repressive chromatin

marks [67]. It has been also previously observed that transgenes inserted into active genes are more resistant to DNA methylation [68] or that endogenous plant genes are more resistant to DNA methylation than transgenes [69].

Interestingly no such difference was observed in the UT lines, where the production/level of sRNAs was even lower than it was for the AS line. The overall amount of 24 nt sRNAs, which are primarily responsible for DNA methylation, was in the UT3 line six-times lower compared to the AS3 line over the region where we analyzed the methylation (the relative proportion of 24 nt sRNAs was similar among the variants; Fig. 5C). This suggests that there might be some qualitative difference between AS and UT sRNAs, which is not obvious from their length or sequence. The likely explanation is that in the UT lines, most of the sRNAs are produced in the nucleus (at least the primary sRNAs). The aberrant transcripts lack a polyA tail, so they are not exported out of the nucleus and must be processed there [41]. In the AS lines, on the other hand, both sense and antisense transcripts could be exported to the cytoplasm, so it is possible that most of the sRNAs were produced in the cytoplasm. Such sRNAs would have a harder time accessing the DNA and it would be more difficult for them to induce DNA methylation. On the other hand, most DCLs seem to be localized in the nucleus [2,70–72] and AGO4 loads sRNAs in the cytoplasm [73], so it remains an open question as to what effect the localization of dsRNA and sRNA production have on their activity. The difference between the ability of AS and UT variants to induce DNA methylation was not caused by the preexisting DNA methylation in the UT lines (Fig. 6A; Fig. S6A). The UT2 line was not methylated before induction and yet it was easily methylated after induction (Fig. 6C; Fig. S6C). Also, we have previously shown that preexisting CG methylation does not enhance methylation by RdDM at all [39].

4 Conclusions

In this work, we compared silencing induced by three different variants of silencing inducers which used different ways to form dsRNAs. Though we supposed that the mechanisms of silencing by the three silencers were well-established, data obtained in our system indicated that in case of the UT silencing, the production of aberrant unterminated RNAs was probably not sufficient to induce silencing, but other features connected with the T-DNA insertions were also required. These might be similar to the S-PTGS that was reported to be induced by read-through antisense transcripts [49].

The three silencers we tested varied considerably regarding silencing dynamics, reversibility, transitivity, and accompanying DNA methylation. The induction of silencing was fastest and strongest with the IR silencer. The AS silencing was most dependent on the induction level. The silencing was maintained after inducer removal specifically in some UT lines. Small RNAs reached the highest levels in the IR variant and lowest in the UT one. Both the AS and UT lines were able to induce strong production of transitive sRNAs, unlike the IR. Surprisingly, in the AS variant, sRNAs showed strongly deviated sRNA-strand ratio shifted in favor of antisense orientation, but the mechanism of it remains unclear.

Our results on DNA methylation accompanying PTGS suggested that there are several factors influencing this phenomenon; i) the strength of induction likely connected with the amount of sRNAs, because IR-induced methylation was strongly dependent on the inducer level, ii) the origin of sRNAs, because sRNAs were present in several times higher levels in the AS variant compared to the UT variant, but the GFP gene was methylated only in the UT variant, iii) the state of the locus targeted by sRNAs, because the silencer GFP was methylated, but the target GFP was not methylated in the AS lines and in the low inducted IR lines despite the sequence identity of these loci. We assume that this sensitivity was affected by the epigenetic state of the locus (e.g. presence of certain chromatin marks), but further research is needed.

Even though we targeted the same sequence at the same locus with the silencing inducers, all of which were based on the same whole *GFP* coding sequence, the silencing responses were surprisingly variable. These differences in behaviors were likely connected with the varying mode of dsRNA precursor formation. Thus, our results raise several questions about specific aspects of RNAi that still await answers.

5 Materials and Methods

5.1 Plant material

Tobacco cell line BY-2 (*Nicotiana tabacum* L. cv. Bright Yellow 2; [36]) was cultivated on modified MS medium with 3% sucrose [74] in darkness at 26 °C. The calli were grown on solidified medium (0.8 w/v agar; 6 cm diameter plates) and subcultured monthly by small pieces (3-5 mm diameter). The suspension cell cultures were cultured in 30 ml of medium in 100 ml Erlenmeyer flasks on an IKA KS501 orbital shaker at 110 rpm (IKA Labortechnik, Staufen, Germany; orbital diameter 30 mm).

The suspension cell cultures were kept in the exponential phase for all experiments by subcultivation every 3-4 days with 2-3 ml of the cell culture (i.e. about 100 mg of fresh cell biomass). Transformations of BY-2 suspension cells were carried out as described previously [38] using *Agrobacterium tumefaciens* strain C58C1 carrying helper plasmid *pGV2260* and an appropriate binary vector (Fig.1; for details on construct preparation see [38]). After cocultivation with agrobacterium, the cells were plated on solidified medium containing 25 μg/ml hygromycin and 100 μg/ml cefotaxim and cultured for 3 weeks. Using this procedure, individual transformed cells formed isolated macroscopic cell clusters (commonly called calli) that could be mostly regarded as genetically homogeneous clones [37].

The XVE inducible system [40] was activated by cultivating the calli or cell suspension cultures on media with an addition of 2μM β-estradiol (Sigma-Aldrich Cat. No. E2758) unless otherwise

specified. The β-estradiol was stored as 20 mM solution in DMSO, therefore a corresponding amount of DMSO was added to the cultivation medium of controls. To remove the β-estradiol from the cell culture (on day 14 of the experiments), 3 ml of the cell culture were washed (diluted and filtered) three times with 50 ml of 3% (w/v) sucrose and then resuspended in a fresh MS medium (30 ml).

5.2 Fluorescence analysis

For population analysis of BY-2 calli, seven calli per plate were grown for 14 days. For each line, one callus was cultivated on medium with estradiol (induced) and one was cultivated on medium without estradiol (mock-induced). Fluorescence was measured every 2-3 days. Each plate was separately photodocumented using G:BOX (SynGene, Cambridge, UK) with a blue excitation light (LED diodes with maximum at λ=465 nm) and a green emission filter (FILT525/GX; 510-540 nm). The images were processed using software NIS-Elements 3.10 (Laboratory Imaging, Prague, CZ). The average light intensity was measured for all the pixels from each callus. Lines that strongly differed in callus fluorescence between the induced and mock-induced variants already at time 0, i.e. lines whose fluorescence did not match the formula: $|\log(|\text{ind.}/\text{mock-ind.}|)| < |\text{average}(\log(\text{ind.}/\text{mock-ind.}))| + 2\times\text{stdev}$ due to non-homogeneous character of the source callus, were excluded from further analyses. Background fluorescence (measured from WT calli) was subtracted from the fluorescence of each callus. After that, the fluorescence of each line was normalized to its mock-induced variant: (induced/mock-induced)×100. Statistical analyses were done by Wilcoxon signed-rank test and Wilcoxon rank-sum test in R 3.4.4.

Fluorescence from cell suspension cultures was measured by flow cytometry. First, protoplasts were prepared by incubating 50-100 mg of cells in 1.5 ml of protoplast enzyme solution: 0.5% (w/v) cellulase R-10 (Duchefa), 0.1% (w/v) pectolyase Y-23 (Duchefa) in 0.45 M mannitol (Sigma-Aldrich) for 3h at 26 °C with gentle shaking. Protoplasts were centrifuged (200 × g for 5 min) and the pellet was gently resuspended in MS medium with 0.4 sucrose. The samples were then measured on LSR II flow cytometer (BD biosciences). On average 14000 cells were measured per

sample. The data were processed by Flowing Software (http://flowingsoftware.btk.fi/) with filtering live protoplasts as described in [75].

5.3 Transcript analyses

RNA was extracted from 100 mg of cells using the phenol-chloroform extraction method [76]. The integrity of the RNA for each sample was checked by gel electrophoresis. For cDNA preparation, 1 μg of the total RNA was treated with DNase I and half of the reaction mixture was then used as a template for RevertAid Reverse Transcriptase (Fermentas, Thermo Fisher Scientific). The other half was subjected to the same treatment without the addition of the reverse transcriptase to check for DNA contamination. A mixture of specific primers was used for the reverse transcription (see Table S1). The final reaction was diluted to 60 μl (template).

The qPCR was done using LightCycler 480 (Roche) and iQ™ SYBR® Green Supermix (BioRad, Hercules, CA, USA) in triplicates of 10 μl reactions with 1 μl of template. We adhered to the qPCR guidelines throughout the experiments [77]. The specificity of the qPCR was verified by melting curve analysis (using the LightCycler 480 software) and also by checking randomly selected samples using gel electrophoresis. The resulting data were processed by LinRegPCR 2015.3 [78], allowing for the correction of the amplicon amplification efficiency. Calculated transcript concentrations were normalized to the *NtEF1α* transcript level, so all the presented values show the relative level of a given transcript to the level of *NtEF1α*. Primers for qPCR are listed in Table S1. Some primer sequences were taken from previous studies [38,79].

5.4 DNA methylation analyses

Bisulfite conversion: The bisulfite conversion was done as described previously [80]. In brief: DNA was extracted using DNeasy Plant Mini Kit (Qiagen) from 100 mg of cells and digested with *Eco*RI. The bisulfite conversion of 600 ng of purified DNA was performed according to the EpiTect Bisulfite Kit (Qiagen). Primers designed to anneal to bisulfite modified DNA (Table S1) were used via PCR

to amplify the regions of interest. The PCR products were cloned into *pDrive* vector (QIAGEN PCR Cloning Kit) and 8-12 clones per sample were sequenced. The resulting sequences were then processed using spreadsheet formulas in LibreOffice 6.1 and the statistical analysis was done using the Wilcoxon rank-sum test in R 3.4.4.

The McrBC assay: The level of methylation was further confirmed by qPCR after cleavage of genomic DNA with McrBC endonuclease (New England Biolabs) specific for methylated DNA (modified from [81]). In brief: 100 ng of DNA was first fragmented with *Ase*I that does not cut in the regions of interest and then the reaction was split in half and supplemented with 10 units of McrBC enzyme or an equivalent amount of 50% (v/v) glycerol. The DNA was digested for 6 h at 37°C, and then the enzyme was inactivated (20 min, 65°C). One ng was used for qPCR performed as described above with primers listed in Table S1.

5.5 Small RNA analyses

Total RNA was extracted from 100 mg (fresh weight) of cells using RNeasy Plant Mini Kit (Qiagen), quality assessed and quantified. A fraction of sRNAs ranging in size from 18–45 nt were excised and recovered from 15% urea-polyacrylamide gels. Extracted sRNAs were ligated with 5′ and 3′ RNA adapters with T4 RNA ligase. The adapter-ligated small RNAs were subsequently transcribed into cDNA by Super-Script II Reverse Transcriptase (Invitrogen) and amplified using adaptor-specific primers. The amplified cDNA products were size-purified and circularized (ssDNA circles). This sRNA library was sequenced using the combinatorial probe-anchor synthesis (cPAS) based BGISEQ-500 sequencer (BGI, Shenzhen, China), which was previously shown to provide highly reproducible results comparable with other NGS platforms [82]. Obtained data were filtered using Geneious 9.1.8 (Biomatters, Ltd.) for reads perfectly matching all the *T-DNAs*. When filtering reads for the AS variant, a letter W (A or T) was placed at the position of the mutation in the sequence (404. nucleotide from the start of *GFP* cds, see Results and Fig. S5H). Data were sorted (based on sequence length, strand orientation and identity of 5’ nucleotide) and sRNA coverage maps were generated using tools

available at usegalaxy.eu [83] while allowing HISAT2 [84] to map each read to every matching position. The maps of coverage were further processed using LibreOffice 6.1 and normalized to the overall amount of sRNAs from the given sample. For the purpose of data presentation, most graphs use average coverage calculated from these maps to allow for comparison of sequence elements of different length. Transcription units (transcription start sites and polyadenylation sites) were defined according to literature data for P_{35S} (including P_{IND} that contains the minimal P_{35S}), P_{NOS} and T_{NOS} [85,86], and for T_{3A} PASPA software (http://bmi.xmu.edu.cn/paspa/index.html) was used to predict the polyadenylation site [87].

5.6 Southern blot analysis

The Southern blot hybridization was done as previously described [15] with few modifications. Briefly, the DNA was isolated from 150 mg (fresh weight) of BY-2 calli. Twenty µg of genomic DNA per sample was separately digested by enzymes *Kpn*I, *Nsi*I and *Ase*I (New England Biolabs). The digoxigenin-labeled probes were prepared by PCR with DIG-dUTP (Roche) according to manufacturer's instructions. Parts of *HPT* and *GFP* genes were amplified with primers listed in Table S1. Hybridization and immunodetection with a chemiluminescent substrate CDP-Star (Tropix) was done according to DIG Application Manual (Roche).

The Southern hybridization results were interpreted as follows: tandem *T-DNA* inserted as a direct repeat should give the same 5.1 kbp fragment with all three restriction enzymes and both probes, plus one fragment of unknown size; head-to-head inverted repeat should give 10.0 kbp fragment when digested with *Kpn*I, two fragments of unknown size with *GFP* probe or 7.4 fragment with *HPT* probe for *Nsi*I and two fragments of unknown size when digested with *Ase*I; tail-to-tail inverted repeat should give two fragments of unknown size with *Kpn*I, 2.8 kbp fragment with *GFP* probe or two fragments of unknown size with *HPT* probe for *Nsi*I and 6.2 kbp fragment when digested with *Ase*I.

6 List of abbreviations

AS: antisense *GFP*
D: DMSO/control
dsRNA: double stranded RNA
E: β-estradiol/induction
GF: control *GFP*
GFP: green fluorescent protein
IR: inverted repeat *GFP*
PTGS: posttranscriptional gene silencing
RNAi: RNA interference
S-PTGS: sense-transgene induced posttranscriptional gene silencing
sRNA: small RNA
TGS: transcriptional gene silencing
UT: unterminated *GFP*

7 Data statement

All data are accessible within the manuscript and its supporting materials. The sRNA-seq datasets generated and/or analyzed during the current study are available in the European Nucleotide Archive, PRJEB33605 (https://www.ebi.ac.uk/ena/data/view/PRJEB33605).

8 Acknowledgments

The work was supported by the Ministry of Education, Youth and Sports of Czech Republic (project No. LO1417).

Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures" (CESNET LM2015042), is greatly appreciated.

We thank Lena Hunt for language corrections.

9 Conflict of interest

The authors declare that they have no competing interests.

10 Author contributions

VC conducted most of the experimental work and participated in all experiments. DT participated in the flow cytometry, RT-qPCR and DNA methylation analyses and also in writing the manuscript. AP participated in methylation analyses. VC together with LF designed the study, interpreted the results and wrote the manuscript. All authors read and approved the final manuscript.

11 References

[1] N.G. Bologna, O. Voinnet, The Diversity, Biogenesis, and Activities of Endogenous Silencing Small RNAs in Arabidopsis, Annu. Rev. Plant Biol. 65 (2014) 473–503. https://doi.org/10.1146/annurev-arplant-050213-035728.

[2] Z. Xie, L.K. Johansen, A.M. Gustafson, K.D. Kasschau, A.D. Lellis, D. Zilberman, S.E. Jacobsen, J.C. Carrington, Genetic and Functional Diversification of Small RNA Pathways in Plants, PLoS Biol. 2 (2004) e104. https://doi.org/10.1371/journal.pbio.0020104.

[3] A. Deleris, J. Gallego-Bartolome, J. Bao, K.D. Kasschau, J.C. Carrington, O. Voinnet, Hierarchical Action and Inhibition of Plant Dicer-Like Proteins in Antiviral Defense, Science. 313 (2006) 68–71. https://doi.org/10.1126/science.1128214.

[4] I.R. Henderson, X. Zhang, C. Lu, L. Johnson, B.C. Meyers, P.J. Green, S.E. Jacobsen, Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning, Nat. Genet. 38 (2006) 721–725. https://doi.org/10.1038/ng1804.

[5] K.D. Kasschau, N. Fahlgren, E.J. Chapman, C.M. Sullivan, J.S. Cumbie, S.A. Givan, J.C. Carrington, Genome-Wide Profiling and Analysis of Arabidopsis siRNAs, PLoS Biol. 5 (2007) e57. https://doi.org/10.1371/journal.pbio.0050057.

[6] S. Mi, T. Cai, Y. Hu, Y. Chen, E. Hodges, F. Ni, L. Wu, S. Li, H. Zhou, C. Long, S. Chen, G.J. Hannon, Y. Qi, Sorting of Small RNAs into Arabidopsis Argonaute Complexes Is Directed by the 5' Terminal Nucleotide, Cell. 133 (2008) 116–127. https://doi.org/10.1016/j.cell.2008.02.034.

[7] E.R. Havecker, L.M. Wallbridge, T.J. Hardcastle, M.S. Bush, K.A. Kelly, R.M. Dunn, F. Schwach, J.H. Doonan, D.C. Baulcombe, The Arabidopsis RNA-Directed DNA Methylation Argonautes Functionally Diverge Based on Their Expression and Interaction with Target Loci, Plant Cell Online. 22 (2010) 321–334. https://doi.org/10.1105/tpc.109.072199.

[8] Z. Zhang, X. Liu, X. Guo, X.-J. Wang, X. Zhang, Arabidopsis AGO3 predominantly recruits 24-nt small RNAs to regulate epigenetic silencing, Nat. Plants. 2 (2016) nplants201649. https://doi.org/10.1038/nplants.2016.49.

[9] A. Takeda, S. Iwasaki, T. Watanabe, M. Utsumi, Y. Watanabe, The Mechanism Selecting the Guide Strand from Small RNA Duplexes is Different Among Argonaute Proteins, Plant Cell Physiol. 49 (2008) 493–500. https://doi.org/10.1093/pcp/pcn043.

[10] A.L. Eamens, N.A. Smith, S.J. Curtin, M.-B. Wang, P.M. Waterhouse, The Arabidopsis thaliana double-stranded RNA binding protein DRB1 directs guide strand selection from microRNA duplexes, RNA. 15 (2009) 2219–2235. https://doi.org/10.1261/rna.1646909.

[11] L. Yang, G. Wu, R.S. Poethig, Mutations in the GW-repeat protein SUO reveal a developmental function for microRNA-mediated translational repression in Arabidopsis, Proc. Natl. Acad. Sci. 109 (2012) 315–320. https://doi.org/10.1073/pnas.1114673109.

[12] M.A. Matzke, R.A. Mosher, RNA-directed DNA methylation: an epigenetic pathway of increasing complexity, Nat. Rev. Genet. 15 (2014) 394–408. https://doi.org/10.1038/nrg3683.

[13] M.-B. Wang, P.M. Waterhouse, High-efficiency silencing of a β-glucuronidase gene in rice is correlated with repetitive transgene structure but is independent of DNA methylation, Plant Mol. Biol. 43 (2000)

67–82. https://doi.org/10.1023/A:1006490331303.
[14] H. Saze, T. Kakutani, Differentiation of epigenetic modifications between transposons and genes, Curr. Opin. Plant Biol. 14 (2011) 81–87. https://doi.org/10.1016/j.pbi.2010.08.017.
[15] E. Nocarova, Z. Opatrny, L. Fischer, Successive silencing of tandem reporter genes in potato (Solanum tuberosum) over 5 years of vegetative propagation, Ann. Bot. 106 (2010) 565–572. https://doi.org/10.1093/aob/mcq153.
[16] T. Pélissier, S. Thalmeir, D. Kempe, H.-L. Sänger, M. Wassenegger, Heavy de novo methylation at symmetrical and non-symmetrical sites is a hallmark of RNA-directed DNA methylation, Nucleic Acids Res. 27 (1999) 1625–1634. https://doi.org/10.1093/nar/27.7.1625.
[17] R. Lister, R.C. O'Malley, J. Tonti-Filippini, B.D. Gregory, C.C. Berry, A.H. Millar, J.R. Ecker, Highly Integrated Single-Base Resolution Maps of the Epigenome in Arabidopsis, Cell. 133 (2008) 523–536. https://doi.org/10.1016/j.cell.2008.03.029.
[18] H.R. Woo, T.A. Dittmer, E.J. Richards, Three SRA-Domain Methylcytosine-Binding Proteins Cooperate to Maintain Global CpG Methylation and Epigenetic Silencing in Arabidopsis, PLOS Genet. 4 (2008) e1000156. https://doi.org/10.1371/journal.pgen.1000156.
[19] A.M. Lindroth, X. Cao, J.P. Jackson, D. Zilberman, C.M. McCallum, S. Henikoff, S.E. Jacobsen, Requirement of CHROMOMETHYLASE3 for Maintenance of CpXpG Methylation, Science. 292 (2001) 2077–2080. https://doi.org/10.1126/science.1059745.
[20] A. Zemach, M.Y. Kim, P.-H. Hsieh, D. Coleman-Derr, L. Eshed-Williams, K. Thao, S.L. Harmer, D. Zilberman, The Arabidopsis Nucleosome Remodeler DDM1 Allows DNA Methyltransferases to Access H1-Containing Heterochromatin, Cell. 153 (2013) 193–205. https://doi.org/10.1016/j.cell.2013.02.033.
[21] C. Liu, F. Lu, X. Cui, X. Cao, Histone Methylation in Higher Plants, Annu. Rev. Plant Biol. 61 (2010) 395–420. https://doi.org/10.1146/annurev.arplant.043008.091939.
[22] J. Du, L.M. Johnson, M. Groth, S. Feng, C.J. Hale, S. Li, A.A. Vashisht, J. Gallego-Bartolome, J.A. Wohlschlegel, D.J. Patel, S.E. Jacobsen, Mechanism of DNA Methylation-Directed Histone Methylation by KRYPTONITE, Mol. Cell. 55 (2014) 495–504. https://doi.org/10.1016/j.molcel.2014.06.009.
[23] X. Li, C.J. Harris, Z. Zhong, W. Chen, R. Liu, B. Jia, Z. Wang, S. Li, S.E. Jacobsen, J. Du, Mechanistic insights into plant SUVH family H3K9 methyltransferases and their binding to context-biased non-CG DNA methylation, Proc. Natl. Acad. Sci. (2018) 201809841. https://doi.org/10.1073/pnas.1809841115.
[24] J.A. Law, J. Du, C.J. Hale, S. Feng, K. Krajewski, A.M.S. Palanca, B.D. Strahl, D.J. Patel, S.E. Jacobsen, Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1, Nature. 498 (2013) 385–389. https://doi.org/10.1038/nature12178.
[25] L.M. Johnson, J. Du, C.J. Hale, S. Bischof, S. Feng, R.K. Chodavarapu, X. Zhong, G. Marson, M. Pellegrini, D.J. Segal, D.J. Patel, S.E. Jacobsen, SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation, Nature. 507 (2014) 124–128. https://doi.org/10.1038/nature12931.
[26] D. Cuerda-Gil, R.K. Slotkin, Non-canonical RNA-directed DNA methylation, Nat. Plants. 2 (2016) 16163. https://doi.org/10.1038/nplants.2016.163.
[27] O. Borsani, J. Zhu, P.E. Verslues, R. Sunkar, J.-K. Zhu, Endogenous siRNAs Derived from a Pair of Natural cis-Antisense Transcripts Regulate Salt Tolerance in Arabidopsis, Cell. 123 (2005) 1279–1291. https://doi.org/10.1016/j.cell.2005.11.035.
[28] V. Ambros, B. Bartel, D.P. Bartel, C.B. Burge, J.C. Carrington, X. Chen, G. Dreyfuss, S.R. Eddy, S. Griffiths-Jones, M. Marshall, M. Matzke, G. Ruvkun, T. Tuschl, A uniform system for microRNA annotation, RNA. 9 (2003) 277–279. https://doi.org/10.1261/rna.2183803.
[29] F. Vazquez, H. Vaucheret, R. Rajagopalan, C. Lepers, V. Gasciolli, A.C. Mallory, J.-L. Hilbert, D.P. Bartel, P. Crété, Endogenous trans-Acting siRNAs Regulate the Accumulation of Arabidopsis mRNAs, Mol. Cell. 16 (2004) 69–79. https://doi.org/10.1016/j.molcel.2004.09.028.
[30] M. Ronemus, M.W. Vaughn, R.A. Martienssen, MicroRNA-Targeted and Small Interfering RNA-Mediated mRNA Degradation Is Regulated by Argonaute, Dicer, and RNA-Dependent RNA Polymerase in Arabidopsis, Plant Cell. 18 (2006) 1559–1574. https://doi.org/10.1105/tpc.106.042127.
[31] C. Napoli, C. Lemieux, R. Jorgensen, Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans., Plant Cell. 2 (1990) 279–289. https://doi.org/10.1105/tpc.2.4.279.

[32] Z. Luo, Z. Chen, Improperly terminated, unpolyadenylated mRNA of sense transgenes is targeted by RDR6-mediated RNA silencing in Arabidopsis, Plant Cell. 19 (2007) 943–958. https://doi.org/10.1105/tpc.106.045724.
[33] C.-F. Chuang, E.M. Meyerowitz, Specific and heritable genetic interference by double-stranded RNA in Arabidopsis thaliana, Proc. Natl. Acad. Sci. 97 (2000) 4985–4990. https://doi.org/10.1073/pnas.060034297.
[34] S.V. Wesley, C.A. Helliwell, N.A. Smith, M. Wang, D.T. Rouse, Q. Liu, P.S. Gooding, S.P. Singh, D. Abbott, P.A. Stoutjesdijk, S.P. Robinson, A.P. Gleave, A.G. Green, P.M. Waterhouse, Construct design for efficient, effective and high-throughput gene silencing in plants, Plant J. 27 (2001) 581–590. https://doi.org/10.1046/j.1365-313X.2001.01105.x.
[35] H. Yan, R. Chretien, J. Ye, C.M. Rommens, New Construct Approaches for Efficient Gene Silencing in Plants, Plant Physiol. 141 (2006) 1508–1518. https://doi.org/10.1104/pp.106.082271.
[36] T. Nagata, Y. Nemoto, S. Hasezawa, Tobacco BY-2 cell line as the "HeLa" cell in the cell biology of higher plants, Int. Rev. Cytol. 132 (1992) 1–30. https://doi.org/10.1016/S0074-7696(08)62452-3.
[37] E. Nocarova, L. Fischer, Cloning of transgenic tobacco BY-2 cells; an efficient method to analyse and reduce high natural heterogeneity of transgene expression, BMC Plant Biol. 9 (2009) 44. https://doi.org/10.1186/1471-2229-9-44.
[38] V. Čermák, L. Fischer, Pervasive read-through transcription of T-DNAs is frequent in tobacco BY-2 cells and can effectively induce silencing, BMC Plant Biol. 18 (2018) 252. https://doi.org/10.1186/s12870-018-1482-3.
[39] A. Přibylová, V. Čermák, D. Tyč, L. Fischer, Detailed insight into the dynamics of the initial phases of de novo RNA-directed DNA methylation in plant cells, Epigenetics Chromatin. 12 (2019) 54. https://doi.org/10.1186/s13072-019-0299-0.
[40] J. Zuo, Q. Niu, N. Chua, An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants, Plant J. 24 (2000) 265–273. https://doi.org/10.1046/j.1365-313x.2000.00868.x.
[41] Y. Huang, G.C. Carmichael, Role of polyadenylation in nucleocytoplasmic transport of mRNA., Mol. Cell. Biol. 16 (1996) 1534–1542. https://doi.org/10.1128/mcb.16.4.1534.
[42] C.M. Hommelsheim, L. Frantzeskakis, M. Huang, B. Ülker, PCR amplification of repetitive DNA: a limitation to genome editing technologies and many other applications, Sci. Rep. 4 (2014) 5052. https://doi.org/10.1038/srep05052.
[43] S.J. Nicholson, V. Srivastava, Transgene constructs lacking transcription termination signal induce efficient silencing of endogenous targets in Arabidopsis, Mol. Genet. Genomics. 282 (2009) 319–328. https://doi.org/10.1007/s00438-009-0467-1.
[44] M.A. Akbudak, S.J. Nicholson, V. Srivastava, Suppression of Arabidopsis genes by terminator-less transgene constructs, Plant Biotechnol. Rep. 7 (2013) 415–424. https://doi.org/10.1007/s11816-013-0278-z.
[45] S. Chen, D. Hofius, U. Sonnewald, F. Börnke, Temporal and spatial control of gene silencing in transgenic plants by inducible expression of double-stranded RNA, Plant J. 36 (2003) 731–740. https://doi.org/10.1046/j.1365-313X.2003.01914.x.
[46] C. Lo, N. Wang, E. Lam, Inducible double-stranded RNA expression activates reversible transcript turnover and stable translational suppression of a target gene in transgenic tobacco, FEBS Lett. 579 (2005) 1498–1502. https://doi.org/10.1016/j.febslet.2005.01.062.
[47] A. Wielopolska, H. Townley, I. Moore, P. Waterhouse, C. Helliwell, A high-throughput inducible RNAi vector for plants, Plant Biotechnol. J. 3 (2005) 583–590. https://doi.org/10.1111/j.1467-7652.2005.00149.x.
[48] P. Pisacane, M. Halic, Tailing and degradation of Argonaute-bound small RNAs protect the genome from uncontrolled RNAi, Nat. Commun. 8 (2017) 15332. https://doi.org/10.1038/ncomms15332.
[49] J.-S. Parent, V. Jauvion, N. Bouché, C. Béclin, M. Hachet, M. Zytnicki, H. Vaucheret, Post-transcriptional gene silencing triggered by sense transgenes involves uncapped antisense RNA and differs from silencing intentionally triggered by antisense transgenes, Nucleic Acids Res. 43 (2015) 8464–8475. https://doi.org/10.1093/nar/gkv753.
[50] M. Kasai, M. Koseki, K. Goto, C. Masuta, S. Ishii, R. Hellens, A. Taneda, A. Kanazawa, Coincident

sequence-specific RNA degradation of linked transgenes in the plant genome, Plant Mol. Biol. 78 (2012) 259–273. https://doi.org/10.1007/s11103-011-9863-0.
[51] M.T. Ruiz, O. Voinnet, D.C. Baulcombe, Initiation and Maintenance of Virus-Induced Gene Silencing, Plant Cell. 10 (1998) 937–946. https://doi.org/10.1105/tpc.10.6.937.
[52] O. Voinnet, P. Vain, S. Angell, D.C. Baulcombe, Systemic Spread of Sequence-Specific Transgene RNA Degradation in Plants Is Initiated by Localized Introduction of Ectopic Promoterless DNA, Cell. 95 (1998) 177–187. https://doi.org/10.1016/S0092-8674(00)81749-3.
[53] H. Vaucheret, C. Béclin, M. Fagard, Post-transcriptional gene silencing in plants, J. Cell Sci. 114 (2001) 3083–3091.
[54] C. Zhang, D.W.-K. Ng, J. Lu, Z.J. Chen, Roles of target site location and sequence complementarity in trans-acting siRNA formation in Arabidopsis, Plant J. 69 (2012) 217–226. https://doi.org/10.1111/j.1365-313X.2011.04783.x.
[55] M. Yoshikawa, T. Iki, H. Numa, K. Miyashita, T. Meshi, M. Ishikawa, A Short Open Reading Frame Encompassing the MicroRNA173 Target Site Plays a Role in trans-Acting Small Interfering RNA Biogenesis, Plant Physiol. 171 (2016) 359–368. https://doi.org/10.1104/pp.16.00148.
[56] I. Szádeczky-Kardoss, T. Csorba, A. Auber, A. Schamberger, T. Nyikó, J. Taller, T.I. Orbán, J. Burgyán, D. Silhavy, The nonstop decay and the RNA silencing systems operate cooperatively in plants, Nucleic Acids Res. 46 (2018) 4632–4648. https://doi.org/10.1093/nar/gky279.
[57] A. Molnar, C.W. Melnyk, A. Bassett, T.J. Hardcastle, R. Dunn, D.C. Baulcombe, Small Silencing RNAs in Plants Are Mobile and Direct Epigenetic Modification in Recipient Cells, Science. 328 (2010) 872–875. https://doi.org/10.1126/science.1187959.
[58] E. Dadami, A. Dalakouras, M. Zwiebel, G. Krczal, M. Wassenegger, An endogene-resembling transgene is resistant to DNA methylation and systemic silencing, RNA Biol. 11 (2014) 934–941. https://doi.org/10.4161/rna.29623.
[59] T. Wroblewski, M. Matvienko, U. Piskurewicz, H. Xu, B. Martineau, J. Wong, M. Govindarajulu, A. Kozik, R.W. Michelmore, Distinctive profiles of small RNA couple inverted repeat-induced post-transcriptional gene silencing with endogenous RNA silencing pathways in Arabidopsis, RNA. 20 (2014) 1987–1999. https://doi.org/10.1261/rna.046532.114.
[60] C. Taochy, N.R. Gursanscky, J. Cao, S.J. Fletcher, U. Dressel, N. Mitter, M.R. Tucker, A.M. Koltunow, J.L. Bowman, H. Vaucheret, B.J. Carroll, A genetic screen for impaired systemic RNAi highlights the crucial role of Dicer-like 2, Plant Physiol. (2017) pp.01181.2017. https://doi.org/10.1104/pp.17.01181.
[61] S. Mlotshwa, G.J. Pruss, A. Peragine, M.W. Endres, J. Li, X. Chen, R.S. Poethig, L.H. Bowman, V. Vance, DICER-LIKE2 plays a primary role in transitive silencing of transgenes in Arabidopsis, PLoS ONE. 3 (2008) e1755. https://doi.org/10.1371/journal.pone.0001755.
[62] J.-S. Parent, N. Bouteiller, T. Elmayan, H. Vaucheret, Respective contributions of Arabidopsis DCL2 and DCL4 to RNA silencing, Plant J. 81 (2015) 223–232. https://doi.org/10.1111/tpj.12720.
[63] S. Polydore, M.J. Axtell, Analysis of RDR1/RDR2/RDR6-independent small RNAs in Arabidopsis thaliana improves MIRNA annotations and reveals unexplained types of short interfering RNA loci, Plant J. 94 (2018) 1051–1063. https://doi.org/10.1111/tpj.13919.
[64] A. Dalakouras, A. Lauter, A. Bassler, G. Krczal, M. Wassenegger, Transient expression of intron-containing transgenes generates non-spliced aberrant pre-mRNAs that are processed into siRNAs, Planta. 249 (2019) 457–468. https://doi.org/10.1007/s00425-018-3015-6.
[65] T.J. Hardcastle, S.Y. Müller, D.C. Baulcombe, Towards annotating the plant epigenome: the Arabidopsis thaliana small RNA locus map, Sci. Rep. 8 (2018) 6338. https://doi.org/10.1038/s41598-018-24515-8.
[66] S.K.T. Ooi, C. Qiu, E. Bernstein, K. Li, D. Jia, Z. Yang, H. Erdjument-Bromage, P. Tempst, S.-P. Lin, C.D. Allis, X. Cheng, T.H. Bestor, DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA, Nature. 448 (2007) 714–717. https://doi.org/10.1038/nature05987.
[67] A. Miura, M. Nakamura, S. Inagaki, A. Kobayashi, H. Saze, T. Kakutani, An Arabidopsis jmjC domain protein protects transcribed genes from DNA methylation at CHG sites, EMBO J. 28 (2009) 1078–1086. https://doi.org/10.1038/emboj.2009.59.
[68] U. Fischer, M. Kuhlmann, A. Pecinka, R. Schmidt, M.F. Mette, Local DNA features affect RNA-directed transcriptional gene silencing and DNA methylation, Plant J. 53 (2008) 1–10. https://doi.org/10.1111/j.1365-313X.2007.03311.x.

[69] L. Vermeersch, N. De Winne, J. Nolf, A. Bleys, A. Kovařík, A. Depicker, Transitive RNA silencing signals induce cytosine methylation of a transgenic but not an endogenous target, Plant J. 74 (2013) 867–879. https://doi.org/10.1111/tpj.12172.
[70] A. Hiraguri, R. Itoh, N. Kondo, Y. Nomura, D. Aizawa, Y. Murai, H. Koiwa, M. Seki, K. Shinozaki, T. Fukuhara, Specific interactions between Dicer-like proteins and HYL1/DRB- family dsRNA-binding proteins in Arabidopsis thaliana, Plant Mol. Biol. 57 (2005) 173–188. https://doi.org/10.1007/s11103-004-6853-5.
[71] P. Hoffer, S. Ivashuta, O. Pontes, A. Vitins, C. Pikaard, A. Mroczka, N. Wagner, T. Voelker, Posttranscriptional gene silencing in nuclei, Proc. Natl. Acad. Sci. 108 (2011) 409–414. https://doi.org/10.1073/pnas.1009805108.
[72] N. Pumplin, A. Sarazin, P.E. Jullien, N.G. Bologna, S. Oberlin, O. Voinnet, DNA Methylation Influences the Expression of DICER-LIKE4 Isoforms, Which Encode Proteins of Alternative Localization and Function, Plant Cell. 28 (2016) 2786–2804. https://doi.org/10.1105/tpc.16.00554.
[73] R. Ye, W. Wang, T. Iki, C. Liu, Y. Wu, M. Ishikawa, X. Zhou, Y. Qi, Cytoplasmic assembly and selective nuclear import of Arabidopsis Argonaute4/siRNA complexes, Mol. Cell. 46 (2012) 859–870. https://doi.org/10.1016/j.molcel.2012.04.013.
[74] T. Murashige, F. Skoog, A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures, Physiol. Plant. 15 (1962) 473–497. https://doi.org/10.1111/j.1399-3054.1962.tb08052.x.
[75] P. Klíma, V. Čermák, M. Srba, K. Müller, J. Petrášek, J. Šonka, L. Fischer, Z. Opatrný, Plant Cell Lines in Cell Morphogenesis Research: From Phenotyping to -Omics, in: F. Cvrčková, V. Žárský (Eds.), Plant Cell Morphog. Methods Protoc., Springer New York, New York, NY, 2019: pp. 367–376. https://doi.org/10.1007/978-1-4939-9469-4_25.
[76] J.L. White, J.M. Kaper, A simple method for detection of viral satellite RNAs in small plant tissue samples, J. Virol. Methods. 23 (1989) 83–93. https://doi.org/10.1016/0166-0934(89)90122-5.
[77] S.A. Bustin, V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M.W. Pfaffl, G.L. Shipley, J. Vandesompele, C.T. Wittwer, The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments, Clin. Chem. 55 (2009) 611–622. https://doi.org/10.1373/clinchem.2008.112797.
[78] C. Ramakers, J.M. Ruijter, R.H.L. Deprez, A.F.M. Moorman, Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data, Neurosci. Lett. 339 (2003) 62–66. https://doi.org/10.1016/S0304-3940(02)01423-4.
[79] G.W. Schmidt, S.K. Delaney, Stable internal reference genes for normalization of real-time RT-PCR in tobacco (Nicotiana tabacum) during development and abiotic stress, Mol. Genet. Genomics MGG. 283 (2010) 233–241. https://doi.org/10.1007/s00438-010-0511-1.
[80] D. Tyč, E. Nocarová, L. Sikorová, L. Fischer, 5-Azacytidine mediated reactivation of silenced transgenes in potato (Solanum tuberosum) at the whole plant level, Plant Cell Rep. 36 (2017) 1311–1322. https://doi.org/10.1007/s00299-017-2155-7.
[81] D.M. Bond, D.C. Baulcombe, Epigenetic transitions leading to heritable, RNA-mediated de novo silencing in Arabidopsis thaliana, Proc. Natl. Acad. Sci. (2015) 201413053. https://doi.org/10.1073/pnas.1413053112.
[82] T. Fehlmann, S. Reinheimer, C. Geng, X. Su, S. Drmanac, A. Alexeev, C. Zhang, C. Backes, N. Ludwig, M. Hart, D. An, Z. Zhu, C. Xu, A. Chen, M. Ni, J. Liu, Y. Li, M. Poulter, Y. Li, C. Stähler, R. Drmanac, X. Xu, E. Meese, A. Keller, cPAS-based sequencing on the BGISEQ-500 to explore small non-coding RNAs, Clin. Epigenetics. 8 (2016) 123. https://doi.org/10.1186/s13148-016-0287-1.
[83] E. Afgan, D. Baker, B. Batut, M. van den Beek, D. Bouvier, M. Čech, J. Chilton, D. Clements, N. Coraor, B.A. Grüning, A. Guerler, J. Hillman-Jackson, S. Hiltemann, V. Jalili, H. Rasche, N. Soranzo, J. Goecks, J. Taylor, A. Nekrutenko, D. Blankenberg, The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update, Nucleic Acids Res. 46 (2018) W537–W544. https://doi.org/10.1093/nar/gky379.
[84] D. Kim, B. Langmead, S.L. Salzberg, HISAT: a fast spliced aligner with low memory requirements, Nat. Methods. 12 (2015) 357–360. https://doi.org/10.1038/nmeth.3317.
[85] H. Guilley, R.K. Dudley, G. Jonard, E. Balàzs, K.E. Richards, Transcription of cauliflower mosaic virus DNA: detection of promoter sequences, and characterization of transcripts, Cell. 30 (1982) 763–773.

https://doi.org/10.1016/0092-8674(82)90281-1.
[86] M. Bevan, W.M. Barnes, M.D. Chilton, Structure and transcription of the nopaline synthase gene region of T-DNA., Nucleic Acids Res. 11 (1983) 369–385. https://doi.org/10.1093/nar/11.2.369.
[87] G. Ji, L. Li, Q.Q. Li, X. Wu, J. Fu, G. Chen, X. Wu, PASPA: a web server for mRNA poly(A) site predictions in plants and algae, Bioinforma. Oxf. Engl. 31 (2015) 1671–1673. https://doi.org/10.1093/bioinformatics/btv004.

12 Additional files

Table S1: List of primers used in this study

Primer Name	Abbreviation	Application	Locus	Sequence	Source
EF1a_qRT-PCR_F		qPCR	*NtEF1α*	TGAGATGCACCACGAAGCTC	Schmidt and Delaney, 2010
EF1a_qRT-PCR_R		qPCR, RT	*NtEF1α*	CCAACATTGTCACCAGGAAGTG	Schmidt and Delaney, 2010
rsGFP-F5	F1, AF2	qPCR, semiqPCR	*RS-GFP*	AAAGATCCCAACGAAAAGAGAGAC	this study
TNOS-R	R1, AR1	qPCR, RT, McrBC assay, se	*TNOS*	GAAACTTTATTGCCAAATGTTTGAACG	this study
rsGFP-R0	F2, SR2	qPCR, semiqPCR	*RS-GFP*	TGTGCCCATTAACATCACCATCTA	this study
PAE3A-R	R2	qPCR, RT, McrBC assay	*T3A*	GAAACCGATGATACGGACGAAAG	this study
BT_LZ-R	R3, AR2	qPCR, RT, McrBC assay, se	region near pl	TCTTCGCTATTACGCCAGATCTTA	this study
HPT-F	F3, AF1	qPCR, semiqPCR	*HPT*	ACCGATGGCTGTGTAGAAGT	this study
NPTII-F	F4	qPCR	*NPTII*	ACGATCCTGAGCGACAATATGA	this study
35S_FG1		bisulfite sequencing	*P35S*	GYAAGTAATAGAGATTGGAG	this study
BS_P35SR		bisulfite sequencing	*RS-GFP*	AACATCACCATCTAATTCAAC	this study
BS_GFPF		bisulfite sequencing	*P35S*	RCACAATCCCACTATCCTTC	this study
BS_GFPR_G		bisulfite sequencing	*RS-GFP*	GTGTGGAYAGGTAATGGTTG	this study
BS_IP_FG		bisulfite sequencing	*PIND*	ATGGATATGTATATGGTGGTAATG	this study
BS_T3A_RC		bisulfite sequencing	*T3A*	TTCCAATRCCATAATACTCAAACT	this study
BS_GFPF2_G		bisulfite sequencing	*RS-GFP*	TTGTTGAATTAGATGGTGATGT	this study
BS_NOST_RC		bisulfite sequencing	*TNOS*	CTCTAATCATAAAAACCCATCTCA	this study
rsGFP F4	F5	McrBC assay	*RS-GFP*	AAGCGTTCAACTAGCAGACCA	Čermák and Fischer, 2018
pER8 50	F6, SF2	McrBC assay, semiqPCR	*PIND*	ATGCTCGACTCTAGGATCTTC	Čermák and Fischer, 2018
GFPgenom_F		probe synthesis	*RS-GFP*	AGTGGAGAGGGTGAAGGTGA	this study
GFPgenom_R		probe synthesis	*RS-GFP*	AAAGGGCAGATTGTGTGGAC	this study
HPTsonda_F		probe synthesis	*HPT*	CGTCTGTCGAGAAGTTTCTGA	Čermák and Fischer, 2018
HPTsonda_R		probe synthesis	*HPT*	ACAGTTTGCCAGTGATACACA	Čermák and Fischer, 2018
rsGFP_F3		T-DNA structure detection	*RS-GFP*	GAGACACCCTCGTCAACAGG	Čermák and Fischer, 2018
XVE-R		T-DNA structure detection	*XVE*	GAGCAAGTTAGGAGCAAACAGTAG	this study
IRdet_F		T-DNA structure detection	*IR-GFP*	AGTTCTTCTCCTTTACTCATGTC	this study
IRdet_R		T-DNA structure detection	*IR-GFP*	ATCCAGAATTCGTGATTGAACTC	this study
PnosF	SF1	semiqPCR	*PNOS*	AAGGGAGTCACGTTATGACC	this study
PnosR	SR1	semiqPCR	*PNOS*	GATTGAGAGTGAATATGAGACTC	this study
BS_35S_RG	SR3	semiqPCR	*P35S*	GAAGGATAGTGGGATTGTG	this study

Combination of primers used in experiments		
estimation of RNA levels		
analyzed RNA	cDNA	qPCR
target GFP	R1	F1+R1
AS silencer	R2	F2+R2
UT silencer	R3	F1+R3
IR silencer	R2	F1+R2
HPT	R1	F3+R1
NPTII	R1	F4+R1

estimation of DNA methylation levels	
analyzed region	qPCR
target GFP	F5+R1
AS silencer	F6+F5
UT silencer	F5+R3
IR silencer	F6+F5

Table S1. List of primers used in this study.

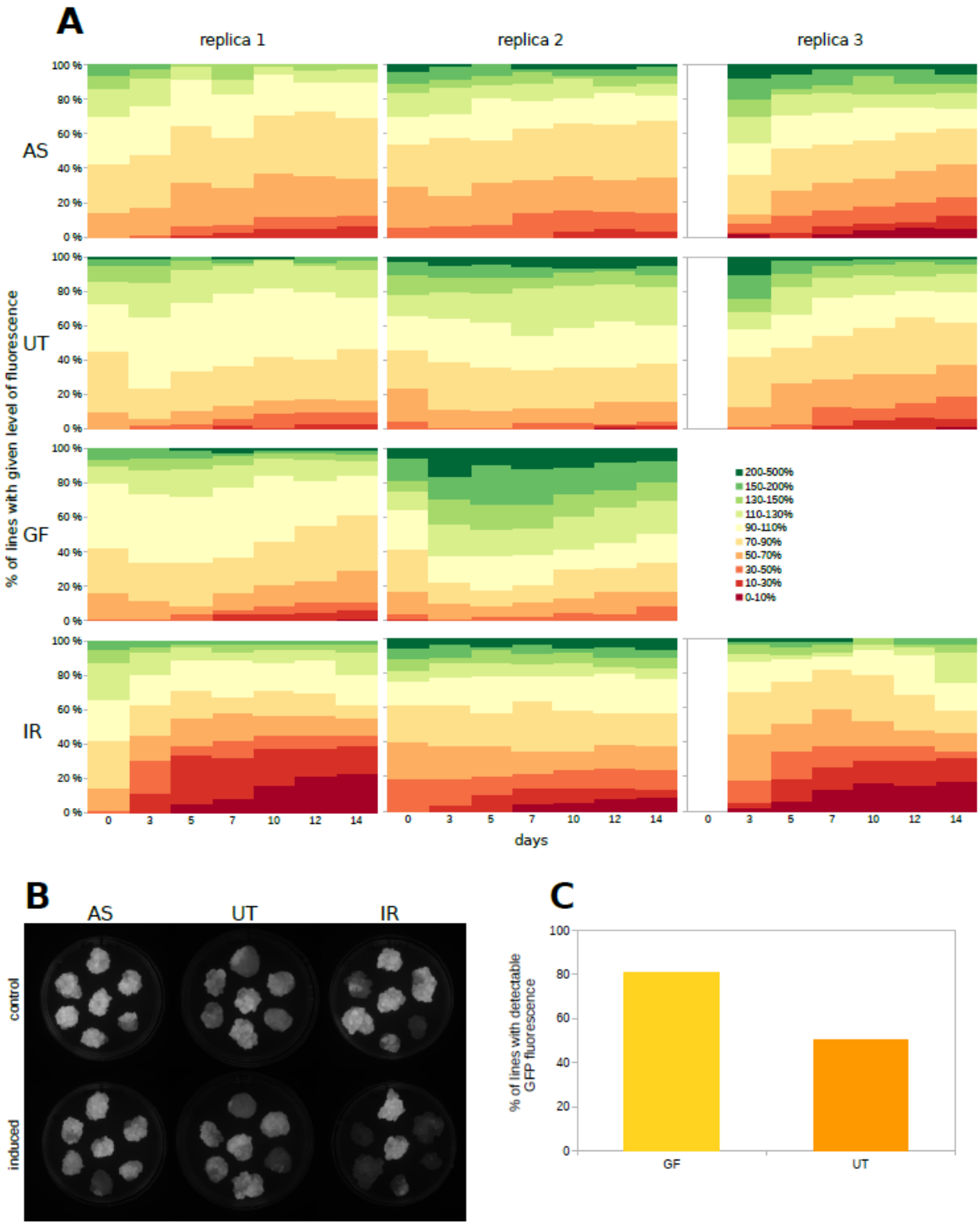


Fig. S1. GFP fluorescence in BY-2 calli carrying various silencer constructs.

(**A**) Changes in the relative representation of lines with a given level of GFP fluorescence during 14-day estradiol treatment (fluorescence levels relative to mock-induced control lines representing 100%). Numbers of analyzed lines for replica #1: AS = 112, UT = 100, GF = 99, IR = 62, for replica #2: AS = 107, UT = 105, GF = 77, IR = 95 and for replica #3: AS = 105, UT = 120, IR = 70. Replicas #2 and #3 were done with UT and GF

constructs that contained *GFP* gene with the start codon (i.e. GFP protein could originate also from the silencer and could influence the results). (**B**) Representative examples of the calli on the 7th day of the experiment. One plate per variant and treatment was chosen. (**C**) Percentage of WT BY-2 lines transformed with UT or GF constructs with the start codon that were able to induce any GFP expression (even when it was only a small patch of a callus mass). Number of analyzed lines was 28 for each variant.

Table S2 A

p-values for comparison of florescence between control and induction medium

days:	0	3	5	7	10
AS	4.1E-2	8.0E-5	3.8E-13	1.3E-8	3.4E-15
UT	1.6E-2	9.9E-1	1.0E-1	1.4E-2	1.5E-4
GF	3.4E-3	6.4E-2	7.9E-2	1.9E-2	8.8E-7
IR	1.7E-1	6.8E-5	1.7E-8	8.2E-8	2.0E-8

$p < 0.001$

Table S2 B

p-values for comparison of rates of silencing between different variants at various time points

day 0	AS	UT	GF	IR
AS		7.6E-1	5.0E-1	8.9E-1
UT			4.5E-1	8.9E-1
GF				4.1E-1
IR				

day 3	AS	UT	GF	IR
AS		6.3E-4	6.6E-2	1.9E-3
UT			1.2E-1	2.1E-6
GF				7.9E-5
IR				

day 5	AS	UT	GF	IR
AS		1.2E-6	7.5E-7	2.0E-3
UT			9.7E-1	1.3E-8
GF				1.2E-8
IR				

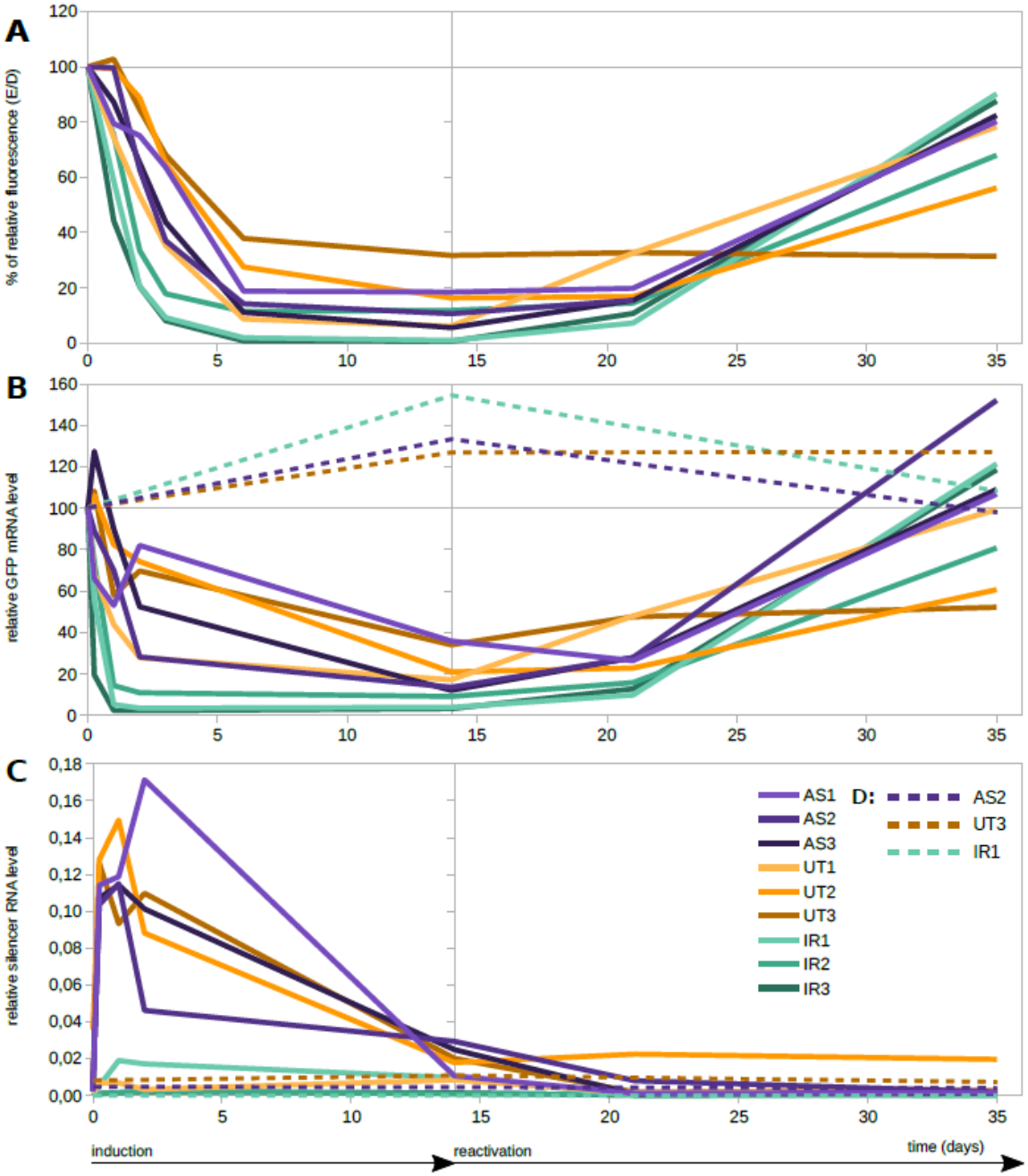
day 7	AS	UT	GF	IR
AS		1.9E-3	5.7E-3	4.6E-4
UT			8.4E-1	6.2E-7
GF				1.9E-6
IR				

day 10	AS	UT	GF	IR
AS		2.0E-5	5.9E-4	6.9E-3
UT			3.0E-1	2.9E-6
GF				2.8E-5
IR				

$p < 0.001$

Table S2. Statistical significance of differences in silencing between silencer variants.

(**A**) p-values for comparison of fluorescence of calli on control and induction medium using Wilcoxon signed-rank test.
(**B**) p-values for comparison of rates of silencing between variants at various time points using Wilcoxon rank-sum test.


A
% of relative fluorescence (E/D)
B
relative GFP mRNA level
C
relative silencer RNA level
AS1
AS2
AS3
UT1
UT2
UT3
IR1
IR2
IR3
D:
AS2
UT3
IR1
induction
reactivation
time (days)

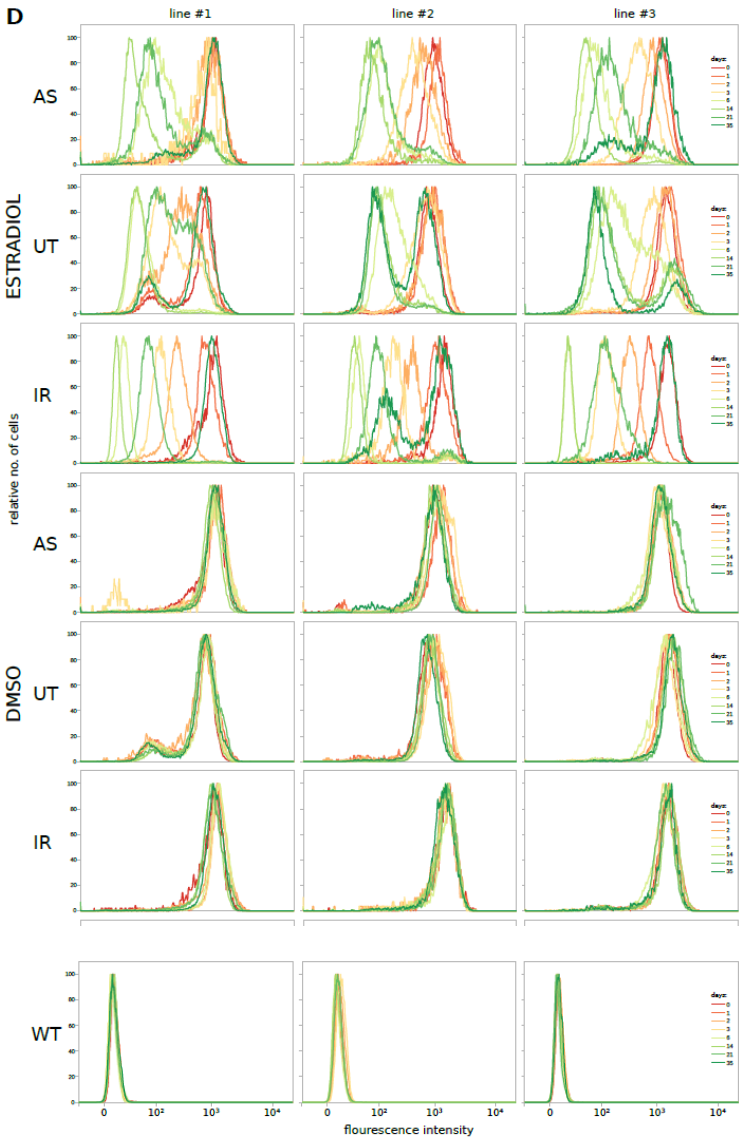
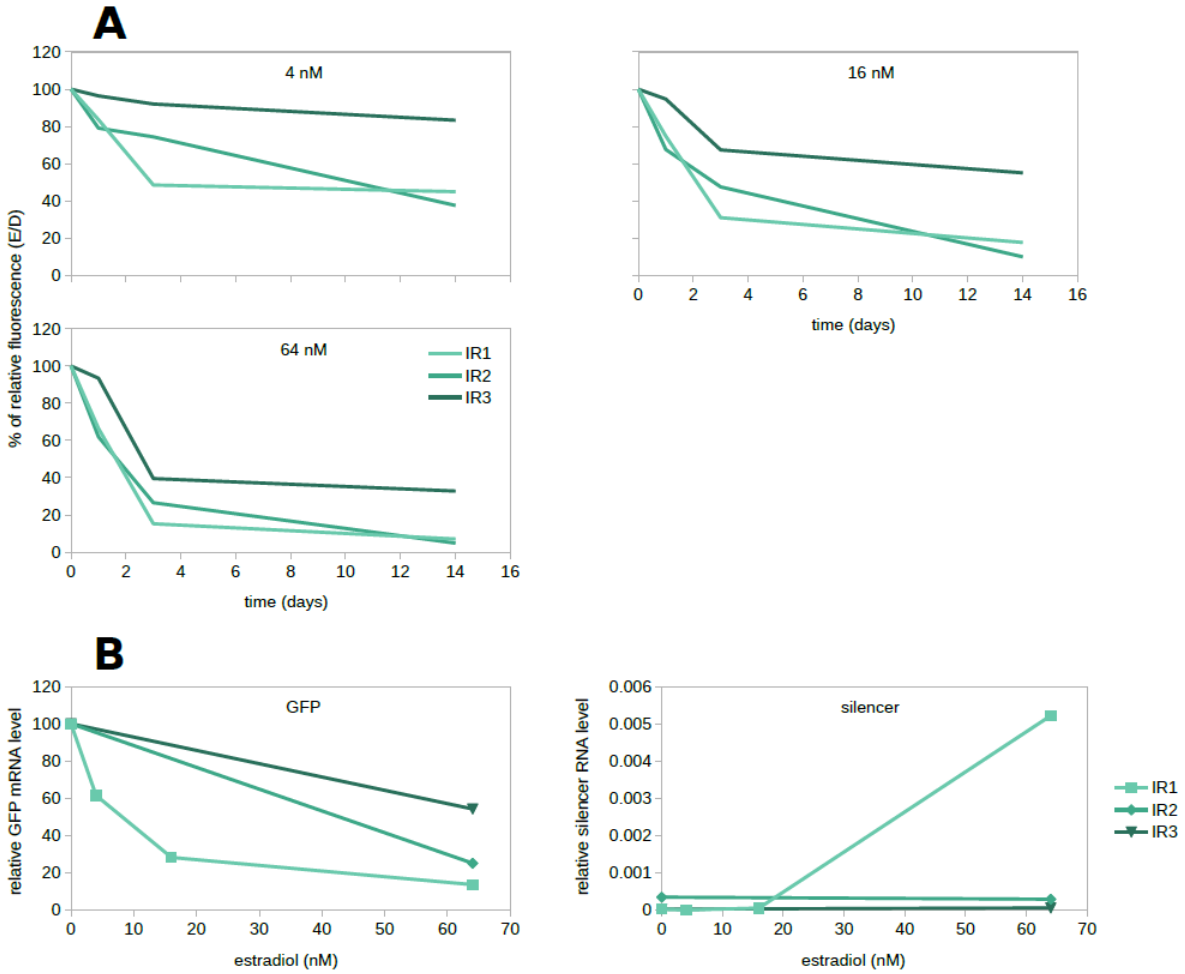


Fig. S2. Dynamics of GFP silencing and recovery in individual selected BY-2 suspension lines.

Three best responding lines for each silencer variant were treated with β–estradiol for 14 days, then β–estradiol was washed out and the cells were monitored for an additional 21 days. (A-C) The same data as presented in Fig. 3A-C but shown individually for each line. In addition, the RNA transcript data also include one line from each variant from the mock-induced control (D). (**A**) Time course of GFP fluorescence as measured by flow cytometer. The relative fluorescence is shown as the ratio of fluorescence between induced and mock-induced cells (E/D). Data were normalized for time 0 to be 100%. (Data for AS2 on the 35th day are not available.) (**B**) Time course of *GFP* transcript levels as measured by RT-qPCR. The relative *GFP* transcript levels were normalized to the internal standard *EF1α* and related to time 0 that was set to 100%. (**C**) Time course of silencer transcript levels as measured by RT-qPCR. The relative silencer transcript levels were normalized to the internal standard *EF1α*. (**D**) Histograms of cell fluorescence levels (flow cytometer data normalized to mode) showing predominantly homogeneous silencing responses in the tested lines.


A
4 nM
16 nM
64 nM
IR1
IR2
IR3
% of relative fluorescence (E/D)
time (days)
B
GFP
relative GFP mRNA level
estradiol (nM)
silencer
relative silencer RNA level
estradiol (nM)

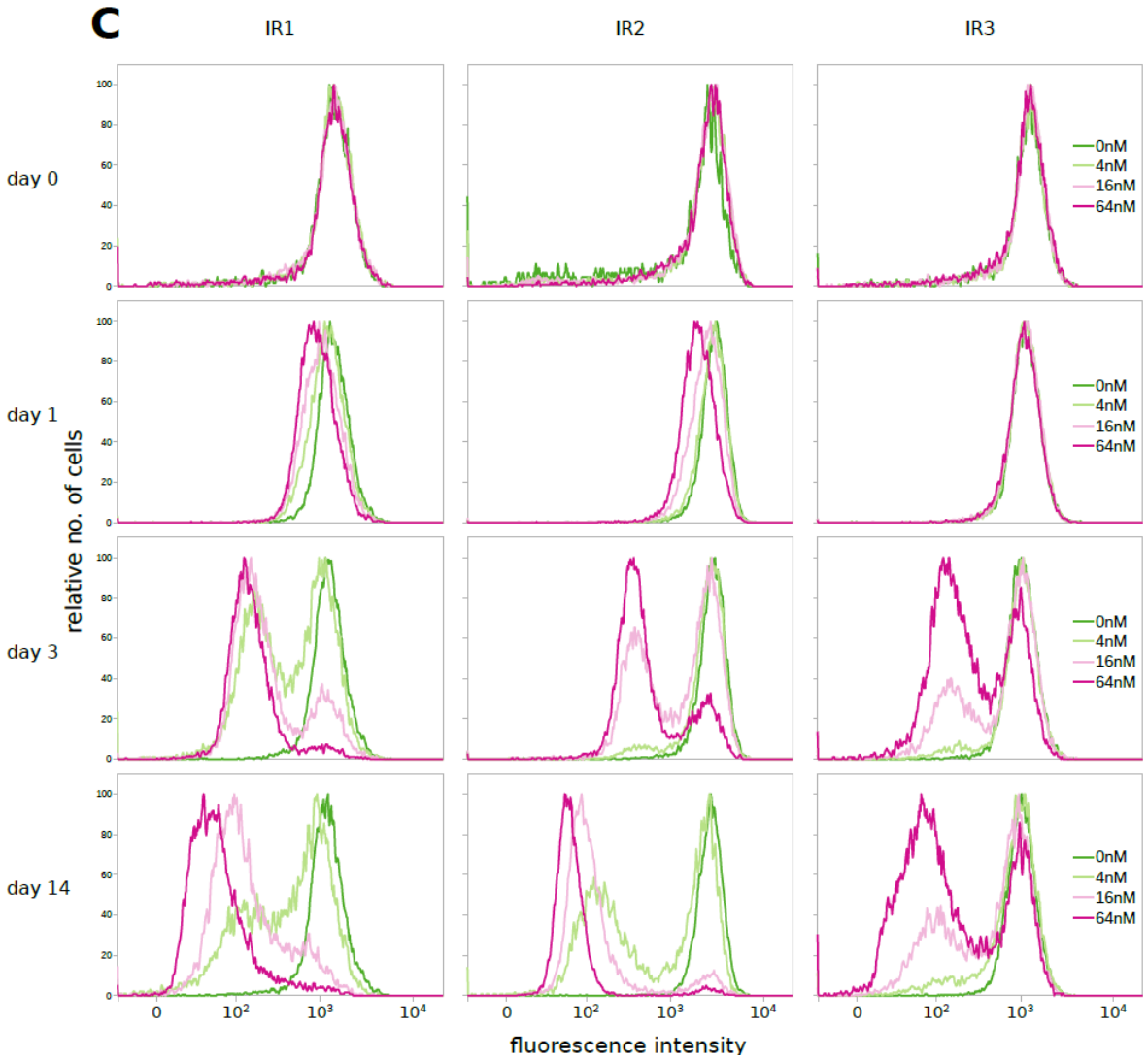
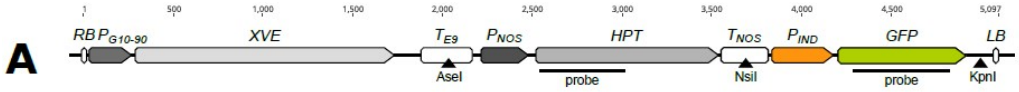
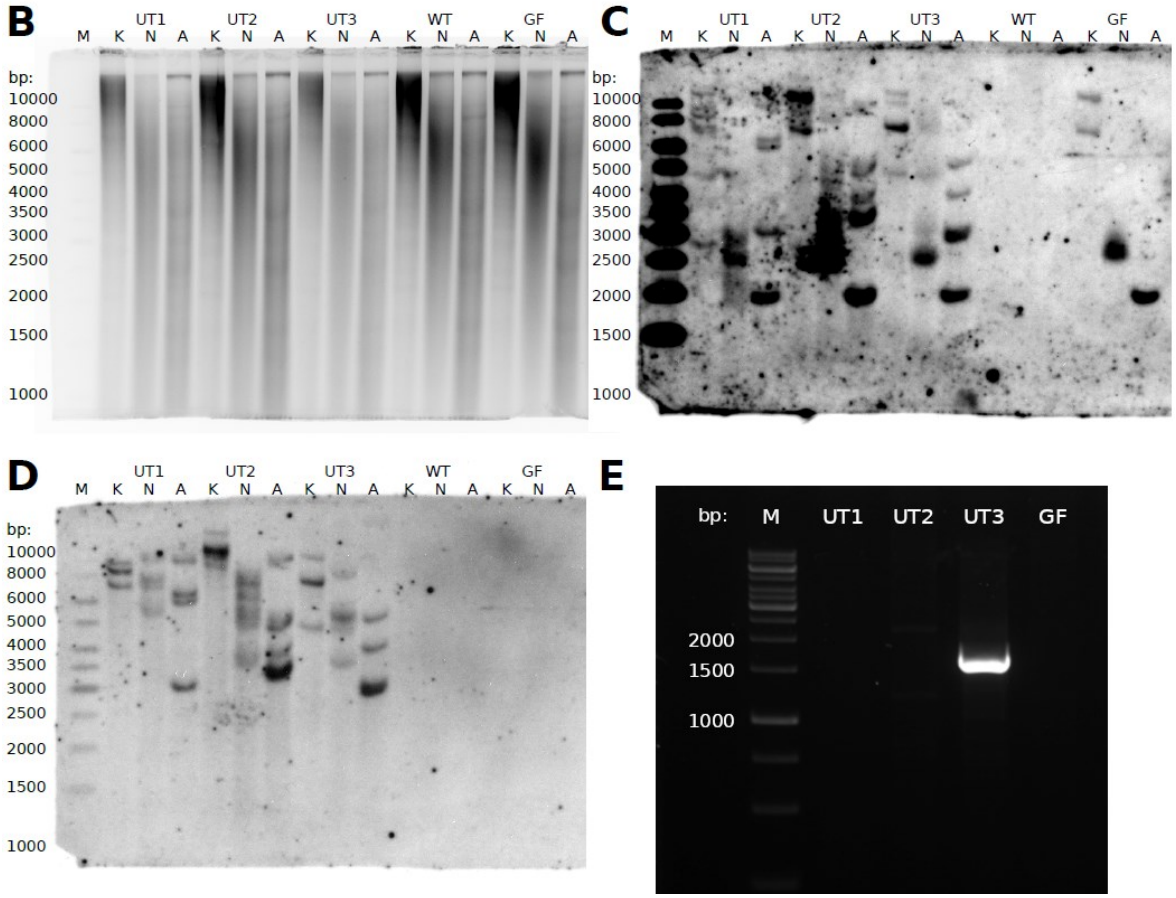


Fig. S3. Effects of estradiol concentration on induction of silencing.

(**A**) Time course of relative GFP fluorescence (as measured by flow cytometer) in individual IR lines (IR1-3) cultured on indicated estradiol concentration. The fluorescence is shown as ratio between the induced and mock-induced suspension cultures (E/D). Data were normalised for time 0 to be 100%. (**B**) Time course of *GFP* and silencer transcript levels as measured by RT-qPCR. The transcript levels were normalized to *EF1α*; for the *GFP* transcript level time 0 was set to 100%; n.d., not determined. (**C**) Histograms of cell GFP fluorescence levels (flow cytometer data normalized to mode) for variants shown in panel A.





F

1: sequence of the T-DNA junction in UT3 line

2: sequence of UT T-DNA tandem repeat junction from complete T-DNAs

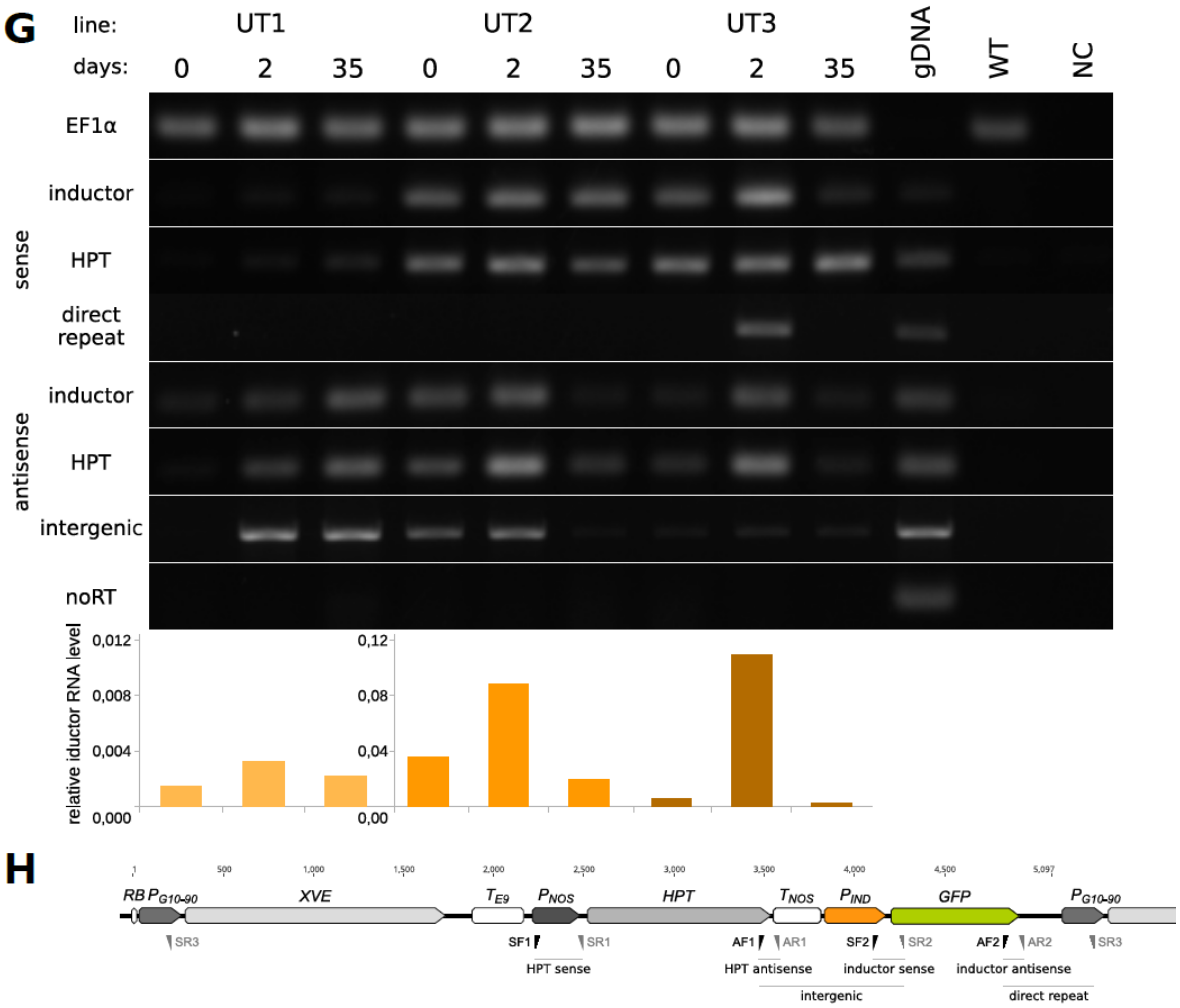


Fig. S4. Detailed characterization of *T-DNA* insertions in UT lines.

(**A**) UT *T-DNA* map with indicated positions of restriction sites and hybridization probes. (**B**) Genomic DNA gel electrophoresis used for Southern blotting. (**C**) Southern blot hybridized with a DIG-labelled *GFP* probe. (**D**) Southern blot hybridized with a DIG-labelled *HPT* probe. (**E**) PCR verification of the direct *T-DNA* tandem repeat predicted in line UT3. The expected length of the PCR product for the perfect direct tandem repeat is 1731bp; PCR primers: *rsGFP-F3* and *XVE-R* (see Table S1). Enzymes used for DNA cleavage: K = *Kpn*I, N = *Nsi*I, A = *Ase*I; M = marker (1kbp ladder, Thermo-Scientific). (**F**) Sequence of tandem *T-DNA* junction in line UT3 aligned to two complete UT *T-DNAs* placed one behind the other (*LB* to *RB*). (**G**) Semiquantitative RT-PCR analyses of read-through transcripts in the UT lines at three different time points. cDNA was prepared with specific primers to distinguish between sense and antisense transcripts. Transcripts at HPT and inducer transcription unit, intergenic region between them and transcript spanning T-DNA junction of direct tandem repeat were tested, EF1α was used as an internal standard, noRT is reaction without reverse transcriptase to ensure that there is no DNA contamination. NC is PCR negative control without template and in gDNA, genomic DNA from line UT3 at a final concentration of 40 pg/µl was used as a template. The graphs of the relative levels of the inducer RNA are from the same data as in Fig. S2C. (**H**) UT *T-DNA* map with indicated positions of primers used for semiquantitative PCR. Primer sequences are listed in Table S1.

Table S3A

			pCP60_GFP		pER8_silencer	
inductor	time (d)	# reads	# mapped	% mapped	# mapped	% mapped
AS	0	49,907,455	1,271	0.00	1,298	0.00
AS	14	41,091,091	226,196	0.55	212,907	0.52
UT	0	37,224,113	8,472	0.02	19,791	0.05
UT	14	42,833,406	103,089	0.24	266,099	0.62
UT	35	45,813,514	138,907	0.30	359,329	0.78
IR	0	43,313,303	44,168	0.10	44,753	0.10
IR	14	48,719,448	901,874	1.85	918,537	1.89
IR	35	60,187,997	3,286	0.01	2,236	0.00

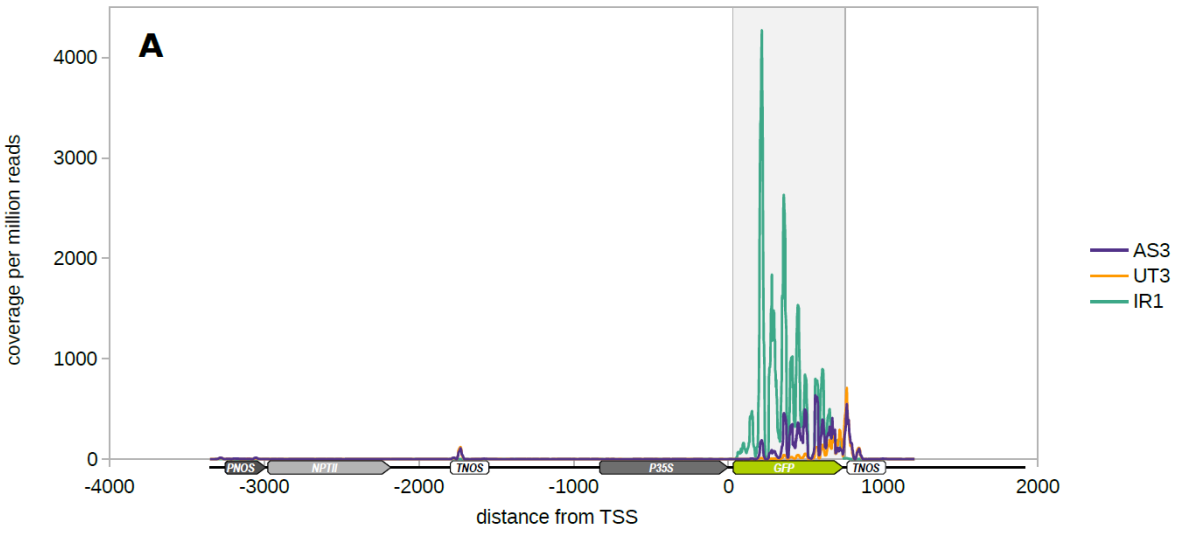
Table S3B

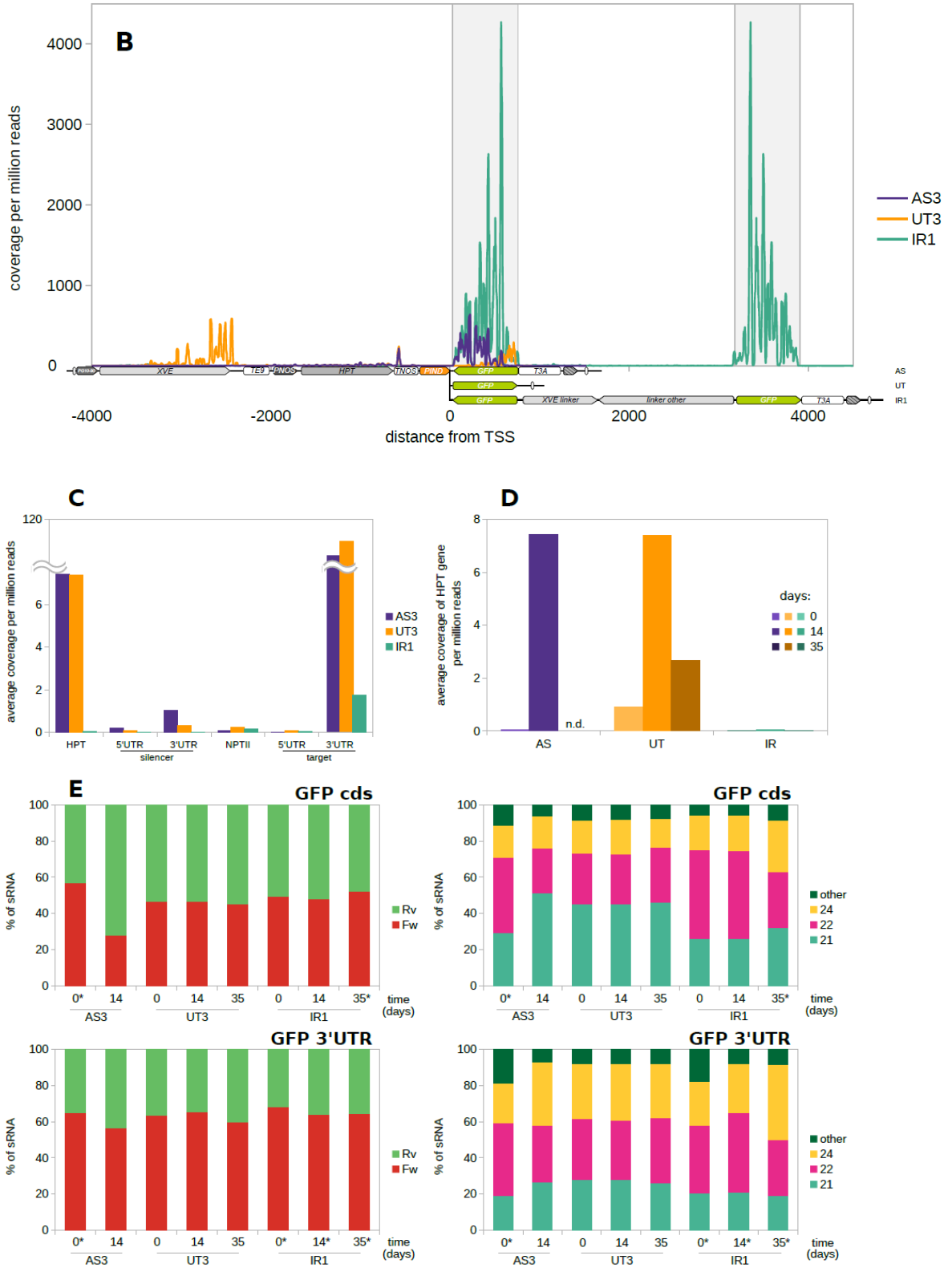
# AS sRNA	Fw	Rv
without mutation	913	5199
with mutation	473	2931

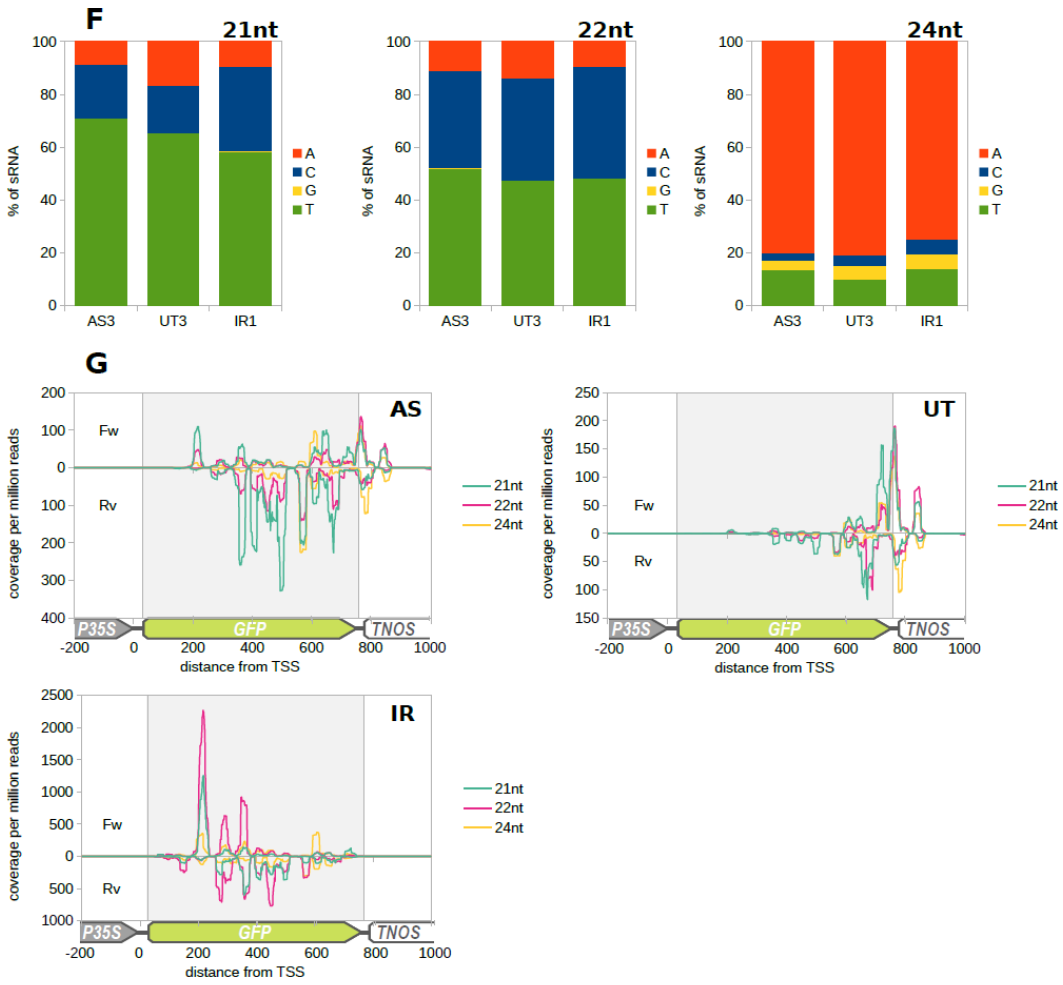
% of sRNA that have to be secondary: 59.6

Table S3. sRNA sequencing statistics.

(**A**) Numbers of reads for each sample and numbers and proportions of reads mapping to the target and silencer *T-DNAs*. (**B**) Numbers of reads mapping to the mutated nucleotide position in the AS silencer. The strand designation "Fw" (forward) and "Rv" (reverse) is relative to the target *GFP*.




B
coverage per million reads
4000
3000
2000
1000
0
-4000
-2000
0
2000
4000
distance from TSS
AS3
UT3
IR1
AS
UT
IR1
XVE
TE9
GNOS
HPT
TNOS
PINO
GFP
T3A
XVE linker
linker other
C
average coverage per million reads
120
6
4
2
0
HPT
5'UTR
3'UTR
NPTII
5'UTR
3'UTR
silencer
target
AS3
UT3
IR1
D
average coverage of HPT gene per million reads
8
6
4
2
0
n.d.
AS
UT
IR
days:
0
14
35
E
GFP cds
% of sRNA
100
80
60
40
20
0
0*
14
0
14
35
0
14
35*
AS3
UT3
IR1
time (days)
Rv
Fw
other
24
22
21
GFP 3'UTR
0*
14*


F
21nt
22nt
24nt
% of sRNA
A
C
G
T
AS3
UT3
IR1
G
AS
UT
IR
coverage per million reads
Fw
Rv
21nt
22nt
24nt
P35S
GFP
TNOS
distance from TSS
H
>pER8-AS-GFP_GFP_sequence
RB
XVE
HPT
GFP
lacZ
LB
I
relative HPT mRNA level
relative NPTII mRNA level
time (days)

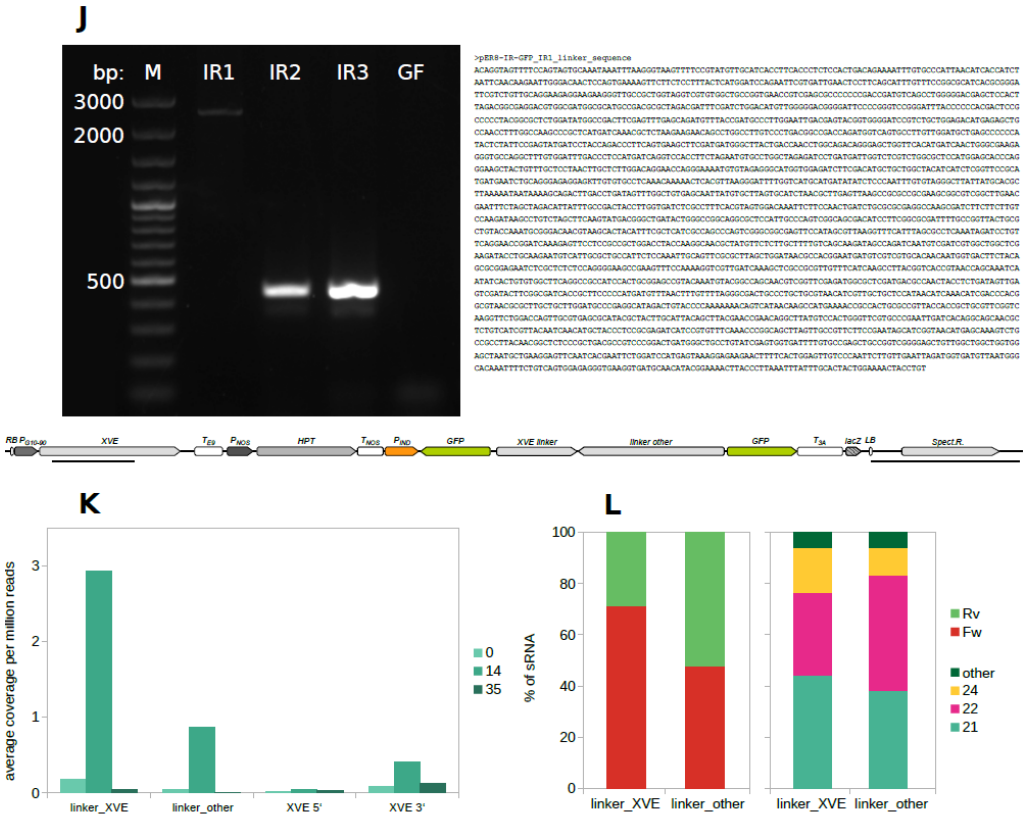
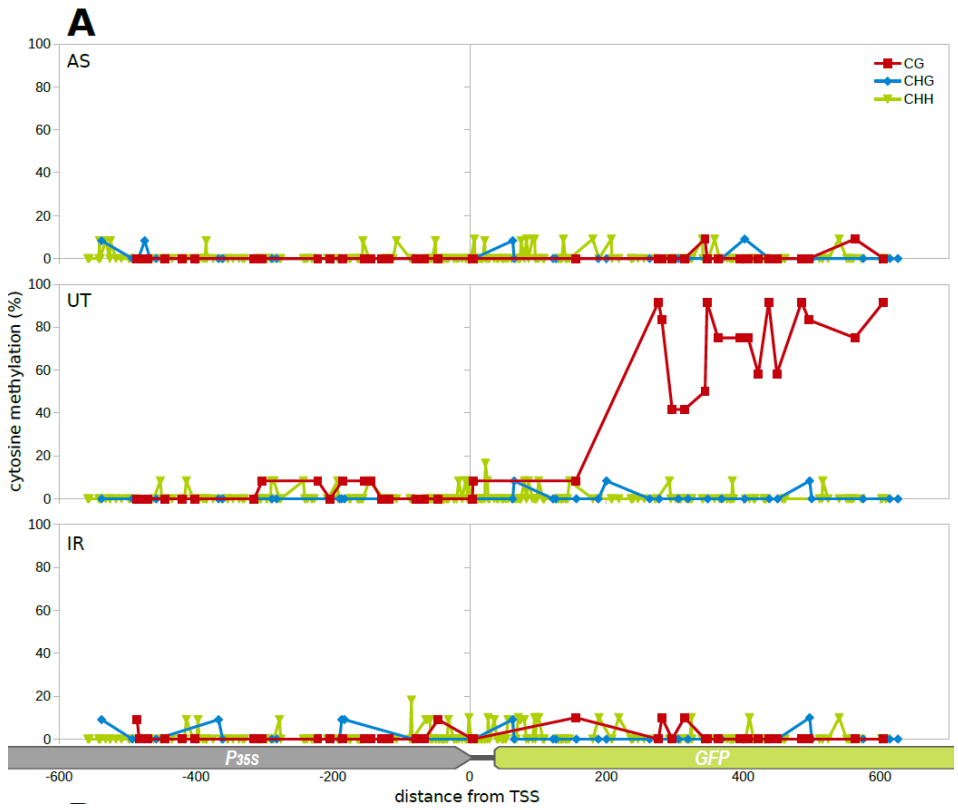
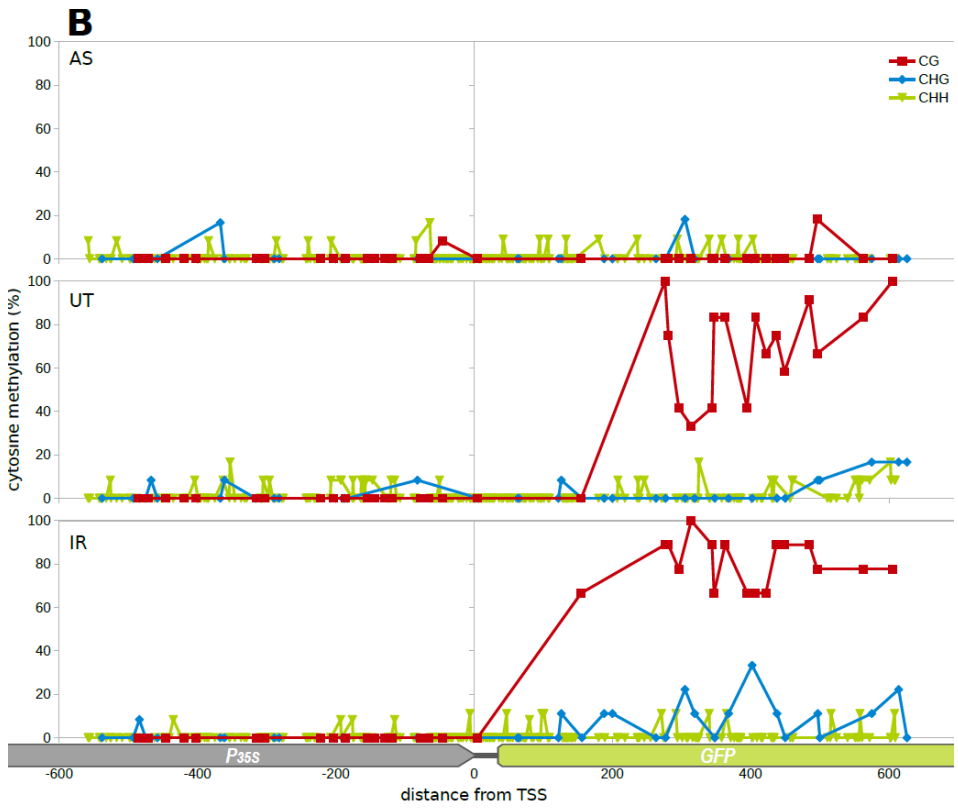


Fig. S5. Characterization of sRNAs matching with the silencer and target *T-DNAs*.

(**A**) Distribution of sRNAs along the whole target *T-DNA* at day 14. (**B**) Distribution of sRNAs along the whole silencer *T-DNA* at day 14. (**C**) Average sRNA coverage of flanking regions of the target and silencer *T-DNA* elements and of coding sequences of resistance genes at day 14. Silencer *5/3'UTR* and target *5/3'UTR* are regions outside the *GFP* coding sequence of the silencer and the target *GFP* transcript, respectively. None of these elements are targeted by the primary sRNAs. (**D**) Average sRNA coverage of the coding sequence of the *HPT* gene for all three lines at different time points. (**E**) Length and strand distribution of sRNAs matching the *GFP* coding region and the terminator *3'UTR* for all three lines at various time points. It should be considered that the statistics for data points marked with asterix (*) are based on very low amounts of sRNAs (see Table S3A). (**F**) Distribution based on the identity of 5'nt for sRNAs of 21, 22 and 24 nt long matching the *GFP* coding sequence at day 14. (**G**) Distribution of sRNAs along the target *GFP* and its *3'UTR* at day 14 based on their length and their strand orientation (forward - Fw are above X axis, reverse - Rv are below X axis). (**H**) Scheme of the silencer *pER8-AS* construct and sequence of the antisense *GFP* with indicated position of the point mutation in red. (**I**) Time course of *HPT* and *NPTII* transcript levels as measured by RT-qPCR. The transcript levels were normalized to *EF1α* and time 0 was set to 100%. (**J**) PCR verification of the linker between inverted repeats in IR1-3 lines. The expected length is 398bp. M = marker (100bp ladder, Thermo-Scientific). PCR primers: *IRdet_F* and *IRdet_R* (see Table S1). The sequence and scheme of the linker region between *GFP* inverted repeats in the IR1 line is also included. Regions of homology between the linker in the IR1 line and the *pER8* plasmid are underlined. (**K**) Average sRNA coverage of the linker region and the *XVE* expression cassette in the IR1 line during different time points. *Linker_XVE* is a part of the linker between inverted *GFP* repeats matching to the transcript of the *XVE* receptor and *linker_other* is the rest of the linker sequence which does not match any sequence in the *T-DNA*, *XVE5'/3'* are parts of the *XVE* gene up and downstream of the *XVE* region present in the linker. (**L**) Length and strand distribution of sRNAs matching to selected elements (as described above) of the target and the silencer *T-DNAs* in the IR1 line at day 14.


A
AS
UT
IR
CG
CHG
CHH
cytosine methylation (%)
P35S
GFP
-600
-400
-200
0
200
400
600
distance from TSS


B
AS
UT
IR
CG
CHG
CHH
cytosine methylation (%)
P35S
GFP
-600
-400
-200
0
200
400
600
distance from TSS

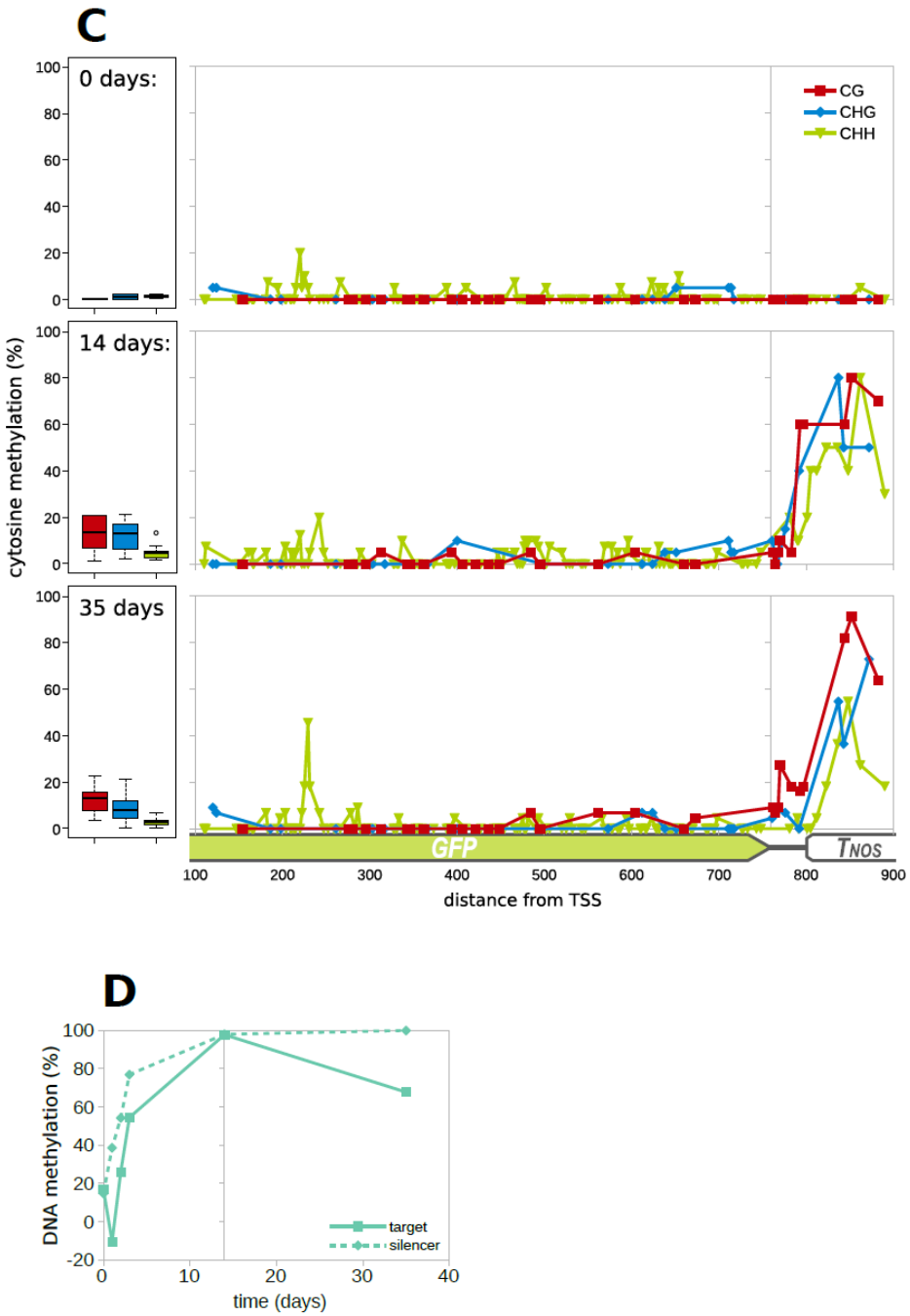


Fig. S6. DNA methylation accompanying PTGS of GFP.

(**A, B**) Distribution of methylated cytosines along the target and 5' adjacent region (A) before the treatment and (B) after 21-day recovery (complementary data to those presented in Fig. 6B). (**C**) Distribution of methylated cytosines in the three sequence contexts along the target region and in the 3' adjacent region for the UT2 line together with boxplots showing average DNA methylation along the whole analyzed region. Analysis was done by bisulfite modification for the three key time points, 10 to 11 DNA clones were sequenced for each data point. (**D**) Dynamics of DNA methylation (assessed by McrBC assay) at the target and silencer *GFP* for line IR1.