Available online at www.sciencedirect.com
SCIENCE DIRECT®
Plant Science 165 (2003) 169–179
PLANT SCIENCE
www.elsevier.com/locate/plantsci

Analysis of the influence of promoter elements and a matrix attachment region on the inter-individual variation of transgene expression in populations of *Arabidopsis thaliana*

Miguel F.C. De Bolle [a,b,1], Katleen M.J. Butaye [a,1], Wim J.W. Coucke [c], Inge J.W.M. Goderis [a,b], Piet F.J. Wouters [a,b], Nadja van Boxel [a], Willem F. Broekaert [a,2], Bruno P.A. Cammue [a,b,*]

[a] *Centrum voor Microbiële en Plantengenetica, Katholieke Universiteit Leuwen, Kasteelpark Arenberg 20, B-3001, Leuwen, Belgium*
[b] *Flanders Interuniversity Institute for Biotechnology-VIB, Kasteelpark Arenberg 20, B-3001, Leuwen, Belgium*
[c] *Laboratorium voor Statistiek en Proeftechniek, Katholieke Universiteit Leuwen, Kasteelpark Arenberg 31, B-3001, Leuwen, Belgium*

Received 10 December 2002; received in revised form 12 March 2003; accepted 13 March 2003

Abstract

Inter-individual variation of transgene expression represents a major bottleneck for high-throughput testing of transgene constructs in plants. More specifically, relatively large populations of first generation transgenic plants are generally required to evaluate transgene constructs with respect to desired expression level and related phenotype. Aiming at a reduction of this inter-transformant variability, in this study we systematically and statistically investigated the influence of different regulatory elements on the level and variability of transgene expression in populations of *Arabidopsis thaliana* plants. Starting from a basic gene construct consisting of the β-glucuronidase reporter gene (*uidA*) controlled by the most commonly used cauliflower mosaic virus (CaMV) 35S promoter, we investigated the effect of a matrix attachment region (MAR), 5′ untranslated regions and the use of different promoter and terminator sequences on the β-glucuronidase expression (GUS expression). It was found that MARs from the chicken lysozyme gene had no influence on the level of expression or on the variability of expression in populations of *A. thaliana* first generation plants. Moreover, transgene expression variability was not influenced by any of four types of terminators nor by any of two different 5′ untranslated regions. Promoters, as expected, drastically influenced expression levels but also rather unexpectedly markedly affected expression variability. A set of promoters derived from the mannopine synthase genes consistently yielded reporter gene expression with higher median levels and lower variance compared to the CaMV 35S promoter. The mannopine synthase promoter derivatives can therefore be useful for a range of applications for which constitutive expression with low inter-individual variability is desired.
© 2003 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: *Arabidopsis thaliana*; Chimeric promoters; MARs; Terminators; Transgene expression; Upstream activators

1. Introduction

Genetic transformation of plants has been widely used over the past 15 years both as a method for studying gene function and as a tool for improving particular traits of plants. It was observed in the early days of plant transformation that the expression level of the transgene can vary considerably between different plants transformed with the same transgene construct [1–3]. This high inter-transformant variability is generally undesirable because it necessitates the screening of large numbers of transformants in order to identify individuals with an acceptable expression level and an expected phenotype. The time and effort spent on screening individual transformants represents a major bottleneck for high-throughput testing of transgene constructs.

* Corresponding author. Fax: +32-16-32-1966.
E-mail address: bruno.cammue@agr.kuleuven.ac.be (B.P.A. Cammue).
[1] Both authors contributed equally.
[2] Present address: CropDesign N.V., Technologiepark 3, B-9052 Gent, Belgium.

0168-9452/03/$ - see front matter © 2003 Elsevier Science Ireland Ltd. All rights reserved.
doi:10.1016/S0168-9452(03)00156-0

Inter-transformant variability of expression levels is thought to be attributable to a number of factors. As the integration of foreign DNA in the plant genome occurs via illegitimate recombination [4], it can take place at virtually any site. As a consequence, sequences flanking the transgene are different in independently obtained transformants. The flanking DNA can contain regulatory elements that may influence transcription of the transgene. Moreover, the chromatin structure of the flanking DNA, which determines accessibility of regulatory sequences by the transcription machinery, can be imposed on the transgene and hence influence transcriptional efficiency [5]. A second source of variation, and probably more important, is the number and configuration of copies of transgenes integrated at one or more loci in the genome [6]. Theoretically, the level of expression should increase with the copy number. However, in most cases studied thus far, expression levels do not correlate with copy number [7,8]. Particular configurations of linked transgene copies like indirect tandem repeats [9] or intrinsic direct repeats [10] are known to cause repeat-induced transgene silencing.

In order to reduce this inter-transformant variability, several research groups have made a number of adaptations to transgene constructs. It has been proposed that, in order to reduce repeat-induced silencing effects, the duplication of sequences within the introduced construct should be avoided [11]. However, the effect of such duplications, for instance identical promoter sequences that drive both the selectable marker gene and the gene of interest, on expression variability has not been systematically analyzed. Most studies have concentrated on the effect of matrix attachment regions (MARs) introduced in vector constructs at positions flanking a gene of interest. MARs are regions in eukaryotic genomes that are attached to the proteinaceous matrix present in the nucleus [12–14]. These regions are generally A/T rich and are typically located in non-transcribed parts of the genome. MARs may act as fixed boundaries of DNA loops showing the same overall chromatin condensation and super helical density. According to the 'loop hypothesis', genes present on such loops are insulated by the MAR boundaries from potential influences of regulatory factors acting on neighboring loops [15]. A number of studies have investigated whether the presence of MARs flanking a transgene results in less variable transgene expression. In most of these studies, MARs significantly increased average transgene expression levels with limited or no decrease of variability in the first generation of transgenic plants [16–21].

One outstanding example of a MAR with strong impact on transgene expression variability, even in the first generation of transgenic plants, is the A element that flanks the chicken lysozyme gene. Presence of this MAR at both ends of the T-DNA borders of an *Agrobacterium tumefaciens*-based transformation vector reduced variance of expression of the *uidA* reporter gene driven by either the potato *Lhca3.St1* gene promoter or the enhanced CaMV 35S promoter in transgenic tobacco plants by 8- and 7-fold, respectively [8,22].

To date, the effect of MARs on expression variability in *Arabidopsis thaliana* has not been thoroughly analyzed. *A. thaliana* is widely used as a model plant for the studies of gene function [23]. Systems allowing low-level variability of transgene expression would facilitate high-throughput analysis of transgenes such as is required for shotgun transformation of bacterial artificial chromosome clones during map-based cloning efforts.

Using *A. thaliana* as a genetic background for large-scale multi-gene delivery experiments, we aim to generate populations of transgenic *A. thaliana* plants that provide minimal inter-transformant expression variability. This study systematically investigates and statistically evaluates the influence of the chicken lysozyme A element MAR and regulatory elements on the level and variability of transgene expression in populations of *A. thaliana* plants, irrespective of T-DNA copy number or integrity of individual transgenic plants.

2. Materials and methods

2.1. Vector construction

For the construction of the different vectors, the modular vector system as fully explained in Goderis et al. [24] was used. Vector pPZP-RCS1 was used to assemble a derivative containing two chicken lysozyme MAR sequences (hereafter abbreviated as chi1MAR). The chi1MAR at the left T-DNA border was inserted as a 3 kb *Xba*I–*BamH*I fragment from plasmid pUC-B-1-X1 [25] in pPZP-RCS1 digested with *Xba*I and *BamH*I. The resulting plasmid was digested with *Kpn*I and *Sal*I and a second copy of chi1MAR was inserted as a 3 kb *Kpn*I–*Sal*I fragment from pUC-B-1-X1 to yield pFAJ3129.

The plasmids pPZP-RCS1 or pFAJ3129 were used to assemble all of the vectors shown schematically in the Figs. 1 and 3. The selectable marker gene cassette was from pMODUL3138 [24]. The reporter gene cassettes were assembled in either pAUX3131 or pAUX 3132 [24] and transferred as I-*Sce*I or I-*Ceu*I fragments, respectively, in the corresponding sites of the plant transformation vectors. The coding region of the reporter gene was taken from the *Escherichia coli* β-glucuronidase gene (*uidA*, as a *Nco*I–*Sac*I-fragment of plasmid pBI505, [26]). The promoters driving the reporter gene were derived from either CaMV 35S ([u*35S*]p*35S*, nucleotides 7019–7342, placed upstream of a segment consisting of nucleotides 7016–7433 [27]) or the man-

Corteva Exhibit 2067 - Page 2 of 11
Inari v. Corteva - PGR2023-00022

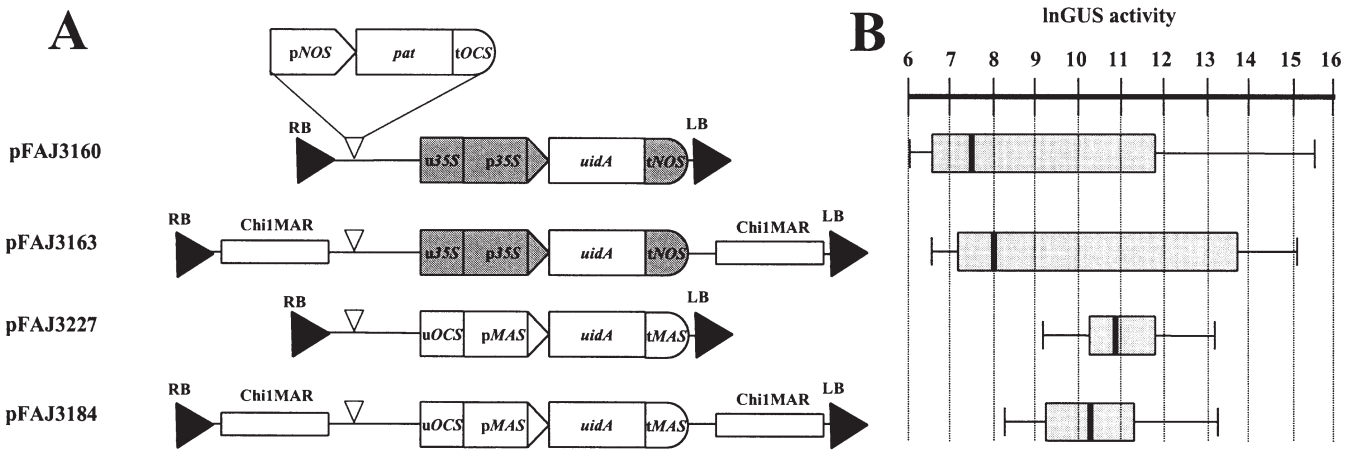

Fig. 1. (A) Schematic representation of the T-DNA regions of transformation vectors pFAJ3160, pFAJ3163, pFAJ3227 and pFAJ3184. Abbreviations are as follows: RB: right T-DNA border; LB: left T-DNA border; Chi1MAR: chicken lysozyme MAR; *uidA*: β-glucuronidase coding region; *pat*: phosphinothricin acetyltransferase coding region; p*NOS*: nopaline synthase promoter; p*35S*: CaMV 35S promoter; p*MAS*: mannopine synthase promoter; t*OCS*: octopine synthase terminator; t*NOS*: nopaline synthase terminator; t*MAS*: mannopine synthase terminator; u*35S*: CaMV upstream activating sequence; u*OCS*: octopine synthase upstream activating sequence. The triangle behind p*35S* represents the tobacco mosaic virus 5′ untranslated sequence; the triangle behind p*MAS* represents the 5′ untranslated sequence of the mannopine synthase gene 2. (B) Box plots of the natural logarithm of the GUS activity (lnGUS) of first generation transgenic *A. thaliana* plants transformed with the vectors pFAJ3160, pFAJ3163, pFAJ3227 and pFAJ3184. LnGUS activity is expressed in pmoles 4-methylumbelliferone per min per mg soluble protein. Gridlines depicted as vertical dotted lines, represent the lnGUS activity scale. Vertical lines in the box plots correspond to the 0, 25th, 50th (or median, indicated in bold), 75th and 100th percentile, respectively. The variance of the lnGUS data is shown by the width of the box plot, whereas the level of lnGUS expression is indicated by the median.

nopine synthase gene (p*MAS*, nucleotides 20128–20633 of the T-DNA region from the *A. tumefaciens* octopine Ti plasmid pTi15955; [28]). Other reporter genes contained any of the following chimeric promoters: [u*OCS*]p*MAS*, [u*OCS*s]p*MAS*s, [u*OCS*]2p*MAS* or [u*OCS*]3p*MAS*. Promoter [u*OCS*]p*MAS* consisted of a fusion between an upstream activating sequence of the octopine synthase promoter (u*OCS*, nucleotides 13775–13991 of the T-DNA region from the *A. tumefaciens* octopine Ti plasmid pTi15955; [28]) and the promoter of the mannopine synthase gene (p*MAS*, nucleotides 20128–20633). Promoter [u*OCS*s]p*MAS*s was identical to [u*OCS*]p*MAS* except that slightly shorter parts of the octopine and mannopine synthase promoters were used (nucleotides 13774–13991 fused to 20128–20513; [28]). Promoters [u*OCS*]2p*MAS* and [u*OCS*]3p*MAS* were identical to [u*OCS*]p*MAS* except that the octopine synthase upstream activating sequence (u*OCS*) was duplicated or triplicated, respectively, as tandem direct repeats. Reporter gene constructs with the [u*35S*]p*35S* promoter had as the 5′ untranslated leader, the 5′ region of the tobacco mosaic virus strain U1 cDNA as described in Ref. [29], while those with p*MAS*, [u*OCS*]p*MAS*, [u*OCS*s]p*MAS*s, [u*OCS*]2p*MAS* and [u*OCS*]3p*MAS* had the 5′ untranslated region of the mannopine synthase gene. A derivative of [u*OCS*]p*MAS* was also made in which the 5′ untranslated region of the mannopine synthase gene (nucleotides 20128–20190 of the T-DNA region from the *A. tumefaciens* octopine Ti plasmid pTi15955; [28]) was replaced with the 5′ untranslated leader region of the tobacco mosaic virus strain U1 cDNA [29]. The terminators in the reporter gene cassette were either from the *A. tumefaciens* nopaline synthase gene (t*NOS*, nucleotides 1298–1551, [30]), the *A. tumefaciens* mannopine synthase gene (t*MAS*, nucleotides 18731–19143 of the T-DNA region from the *A. tumefaciens* octopine Ti plasmid pTi15955; [28]), the *A. tumefaciens* octopine synthase gene (t*OCS*, nucleotides 12266–12556) or the *A. tumefaciens* T-DNA gene 7 (t*G7*, nucleotides 2612–2914; [28]).

2.2. Plant transformation

All plant transformation vectors were introduced by electroporation in *A. tumefaciens* strain GV3101 containing the pGV3850 helper plasmid [31]. The *A. tumefaciens* strains with the binary vectors were used to transform *A. thaliana* ecotype Col-0 plants using the vacuum infiltration method of Bechthold et al. [32]. Transgenic plants were selected as described by Goderis et al. [24] and grown in an environmentally controlled growth chamber with 22 °C day-time temperature, 18 °C night-time temperature and a 12 h photoperiod (photon flux of 100 μE m^{-2} s^{-1}).

Primers for verifying transgenic of *A. thaliana* plants were the following: OWB425 (GGTGGGAAAGCGCGTTACAAG; sense primer annealing to the *uidA* gene); OWB448 (AACAGGATTCAATCTTAAGAAACTT; antisense primer annealing to the t*NOS* sequence); OWB449 (CATAAATTGATGATATAGC-

Corteva Exhibit 2067 - Page 3 of 11
Inari v. Corteva - PGR2023-00022

TAGCTT; antisense primer annealing to the t*G7* sequence); OWB450 (GTTTGTTTGGTTGCCTTTGCAA; antisense primer annealing to the t*MAS* sequence); OWB451 (GCGATCATAGGCGTCTCGCAT; antisense primer annealing to the t*OCS* sequence). Genomic DNA was obtained according to Eggermont et al. [33]. All polymerase chain reactions (PCR) were performed as described by Sambrook et al. [34]. First generation transgenic *A. thaliana* plants were grown until they were 5 weeks old. At this stage the 4th, 6th and 8th leaf of each plant were collected in a single tube, flash frozen in liquid nitrogen and stored in −80 °C prior to GUS activity analysis. Stems and roots were obtained from mature, 8-week-old plants and these tissues were treated similarly.

2.3. Enzyme assays

β-Glucuronidase (GUS) activity was measured fluorometrically using 4-methylumbelliferyl glucuronide as a substrate and 4-methylumbelliferon as a standard according to Jefferson [35]. Total protein was determined by the method of Bradford [36] using bovine serum albumin as a standard.

2.4. Statistical analysis

Statistical analysis was performed with the statistical package R that can be downloaded from the internet at http://www.cran.r-project.org. The freely available software offers an object oriented programming environment including pre-defined functions. Normal distribution of datasets is required to meet the robustness properties of the standard statistical applications. QQ-plots and density plots revealed a non-normal distribution of some of the datasets analyzed, even after natural logarithm transformation [37]. The non-normality of the datasets implies the use of methods, which do not require any assumption about normality [37]. Therefore a bootstrap procedure was performed to analyze our datasets instead of the standard statistical applications. The bootstrap method is based on a resampling procedure and is used for both the median and variance comparison tests [38]. To compare the variances of the datasets a bootstrap based on Bartlett's test for homogeneity of variances was used as described in Ref. [39]. The tests were performed for a significance level of 95%, taking into account the adjustments for multiple comparisons. The *P*-values were adjusted for simultaneous hypotheses as suggested by Westfall and Young [40].

3. Results

3.1. The effect of chicken lysozyme MARs on reporter gene expression

A series of plant transformation vectors (Fig. 1A) was constructed in order to study the effect of the chicken lysozyme MAR sequences on reporter gene expression in transgenic *A. thaliana* plants. Vector pFAJ3163 was constructed such that its T-DNA region has the same configuration as that of pLM9, a vector demonstrated by Mlynárová [22] to provide extremely low inter-transformant variability of gene expression in tobacco due to the presence of chicken lysozyme MAR sequences flanking each T-DNA border. The relative position and orientation of the chicken lysozyme MAR sequences as well as that of the selectable marker and reporter genes is identical in pFAJ3163 and pLM9. In both vectors the reporter gene is composed of the enhanced CaMV 35S promoter ([u*35S*]p*35S*), the *uidA* coding region and the nopaline synthase terminator (t*NOS*). A variant of pFAJ3163, construct pFAJ3184, allowed us to investigate whether or not the effect of the chicken lysozyme MAR is dependent on the nature of the regulatory sequences of the reporter gene. Vector pFAJ3184 is identical to pFAJ3163 except that it has the *uidA* reporter gene controlled by a mannopine synthase terminator (t*MAS*) and a chimeric promoter [u*OCS*]p-*MAS*. The latter consists of an enhancer element from the octopine synthase gene fused to the mannopine synthase promoter. This chimeric promoter is nearly identical to the constitutive promoter A*ocs*A*mas*P*mas* shown by Ni et al. [41] to confer a higher level of expression than that of the enhanced CaMV 35S promoter, at least in transgenic tobacco plants. Vectors corresponding to pFAJ3163 and pFAJ3184 but lacking the chicken lysozyme MARs were also constructed and called pFAJ3160 and pFAJ3227, respectively.

For each of the vectors *A. thaliana* populations consisting of at least 20 primary transformants were obtained and their transgenic nature was assessed by growth in presence of the selective agent (phosphinothricin) and by PCR amplification of a region spanning the terminator and a part of the *uidA* reporter gene (Fig. 2). Bands of the correct size (1800 bp) were unambiguously present in the lanes representing the amplification products of the transformants (lanes 4–6 and 8–10) and plasmid DNA (lanes 1 and 13). GUS activity was determined on a mixture of the 4th, 6th and 8th leaf collected from each plant. In an attempt to obtain normally distributed data sets, the GUS activity values were transformed by taking the natural logarithm (lnGUS), as previously done [8,17,22,37]. The lnGUS data from plants transformed with the constructs pFAJ3227 and pFAJ3184, showed a nearly normal distribution whereas lnGUS data from plants trans-

Corteva Exhibit 2067 - Page 4 of 11
Inari v. Corteva - PGR2023-00022

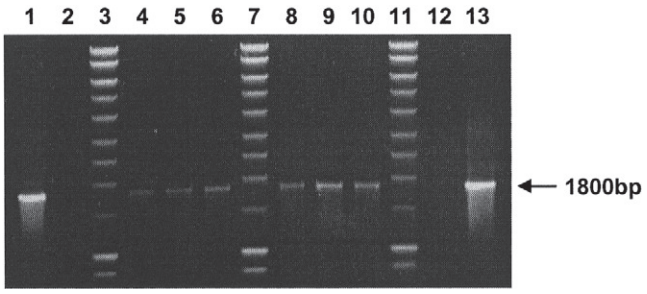

Fig. 2. Analysis of the PCR products amplified from plasmid DNA (lanes 1 and 13), from total genomic DNA of non-transformed plants (lanes 2 and 12) and transgenic plants (lanes 4–6 and 8–10). DNA Smart Ladder marker in lanes 3, 7 and 11. An arrow indicates the expected size in base pairs.

formed with pFAJ3160 and pFAJ3163 were bimodally distributed (Fig. 3). The non-normality of these data sets imposed the use of statistical methods that are not based on normal distribution of the data and are hence based on the analysis of median instead of average expression levels (see Section 2.4).

The lnGUS activity data of the *A. thaliana* plants transformed with pFAJ3160, pFAJ3163, pFAJ3227 or pFAJ3184 are presented as box plots in Fig. 1B and the statistical evaluation of these data is shown in Table 1. Neither the median nor the variance of the lnGUS activity of pFAJ3160 or pFAJ3227 transformed lines were significantly different from those of pFAJ3163 or pFAJ3184 transformed lines, respectively (Table 1). The [u*OCS*]p*MAS*–*uidA*–t*MAS* transcription unit in constructs pFAJ3184 and pFAJ3227 yielded, however, expression levels of which the median was significantly higher than that of the transcription unit [u*35S*]p*35S*–*uidA*–t*NOS* in constructs pFAJ3163 and pFAJ3160, irrespective of the presence or absence of the chicken lysozyme MARs (Table 1). In addition, the variance of lnGUS activity was significantly lower for plants transformed with the [u*OCS*]p*MAS*–*uidA*–t*MAS* transcription unit than for plants transformed with the [u*35S*]p*35S*–*uidA*–t*NOS* transcription unit, again irrespective of the presence or absence of the chicken lysozyme MARs (Table 1).

3.2. The effect of the mannopine synthase promoter on the reporter gene expression

The transcription units [u*35S*]p*35S*–*uidA*–t*NOS* and [u*OCS*]p*MAS*–*uidA*–t*MAS* in the vectors pFAJ3163 and pFAJ3184, respectively, differ in several components: (i) the promoters (p*35S* in pFAJ3163 and p*MAS* in pFAJ3184), (ii) the upstream activating sequences (u*35S* in pFAJ3163 and u*OCS* in pFAJ3184), (iii) the terminators (t*NOS* in pFAJ3163 and t*MAS* in pFAJ3184) and (iv) the 5′ untranslated leader (derived from tobacco mosaic virus in pFAJ3163 and derived from the mannopine synthase gene in pFAJ3184). To

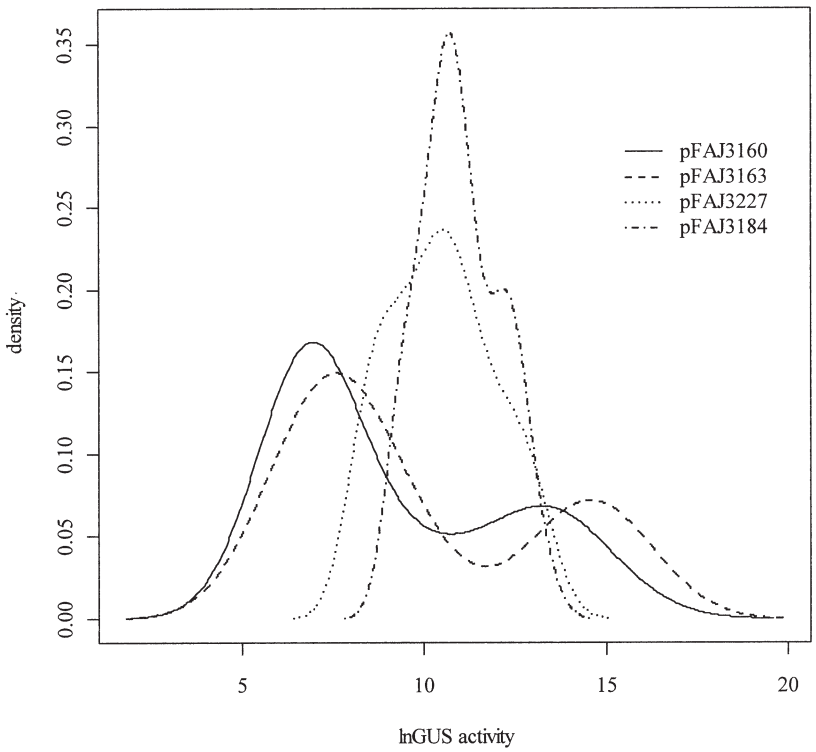

Fig. 3. Distribution plots representing the lnGUS data obtained from transgenic plants transformed with the constructs shown in Fig. 1. Density is expressed as the ratio of the number of observations to the total number of observations for a certain value of lnGUS activity. LnGUS activity is expressed in pmoles 4-methylumbelliferone per min per mg soluble protein.

Table 1
Statistical analysis of expression levels in T0 generation populations of *A. thaliana* plants transformed with the constructs shown in Fig. 1A

Construct	MAR	No[a]	GUS activity[b]		lnGUS activity[c]			
			Median	Mean	Median	P^d	Variance	P^d
pFAJ3160	−	34	1732	424 600	7.46	b	9.26	b
pFAJ3163	+	23	3036	708 670	8.02	b	11.14	b
pFAJ3227	−	32	49 730	100 080	10.81	a	1.15	a
pFAJ3184	+	33	31 265	90 840	10.35	a	2.09	a

[a] Number of plants analyzed in the T0 generation of transformants.
[b] GUS activity is expressed in pmoles 4-methylumbelliferone per min per mg soluble protein.
[c] Natural logarithm of GUS activity.
[d] Comparisons of median or variance at a significance level 0.05. Identical letters indicate that the data (median or variance) are not significantly different.

systematically investigate which of these components is responsible for the decreased inter-transformant expression variability, a series of derivatives of pFAJ3184 was constructed in which one or more of the components were replaced by a component of pFAJ3163. This vector series, depicted in Fig. 4A, was used to generate populations of transgenic *A. thaliana* plants. Data of lnGUS activities of these populations are shown in Fig. 4B as box plots and the statistical analysis is presented in Table 2.

Vector pFAJ3357, differing from pFAJ3184 by the replacement of [u*OCS*]p*MAS* by [u*35S*]p*35S*, yielded a variance of the lnGUS values that was significantly higher than observed for pFAJ3184, being 6.15 for the former versus 2.09 for the latter vector. Moreover, the median lnGUS activity obtained with pFAJ3357-transformed plants (8.59) was significantly lower than that of pFAJ3184-transformed plants (10.35).

The effect of the upstream activating sequence of the octopine synthase promoter (u*OCS*) was assessed in two different ways. Firstly, the u*OCS* element was added upstream of the [u*35S*]p*35S* promoter. Transformation with this construct, pFAJ3293, resulted in significantly higher expression levels (median lnGUS activity of 12.14 versus 10.35) and a significantly higher variance (12.09 versus 2.09) when compared to pFAJ3184 transformed lines. Secondly, deletion of the u*OCS* element from the [u*OCS*]p*MAS* promoter, as in construct pFAJ3183, resulted in a significant reduction of median lnGUS activity (10.35 and 7.80 for pFAJ3184 and pFAJ3183, respectively) and in a small, not significant reduction on the variance of lnGUS activity (2.09 and 1.06 for pFAJ3184 and pFAJ3183, respectively). Addition of one more u*OCS* element to the [u*OCS*]p*MAS* promoter to yield the [u*OCS*]2p*MAS* promoter (construct pFAJ3185), slightly increased the median lnGUS activity while slightly decreasing the variance of the lnGUS activity. However, these effects were not statistically significant. Finally, addition of two more u*OCS* elements to the [u*OCS*]p*MAS* promoter to obtain the [u*OCS*]3p*MAS* promoter (construct pFAJ3186) resulted in a slight but non-significant decrease of lnGUS variance (2.09 versus 1.12 for pFAJ3184 or pFAJ3186 transformed plants, respectively) and a significant increase of median lnGUS activity (12.28 versus 10.35 for pFAJ3186 or pFAJ3184 transformed plants, respectively).

We also tested whether alterations in the 3′ and 5′ untranslated regions would affect expression variability. Vectors pFAJ3184, pFAJ3269, pFAJ3270 and pFAJ3271 which only differed in the terminators of the reporter genes (t*MAS*, t*NOS*, t*G7* and t*OCS*, respectively) generated populations of which neither the median nor the variance of the lnGUS values was significantly different. In addition, populations of plants transformed with vectors pFAJ3184 and pFAJ3385, which only differed at the 5′ untranslated region, were statistically indistinguishable in terms of median and variance of lnGUS activity.

The [u*OCS*]p*MAS* promoter is nearly identical to the promoter designated as A*ocs*A*mas*P*mas* by Ni et al. [41]. The u*OCS* element in [u*OCS*]p*MAS* corresponds to the −339/−116 domain of the octopine synthase promoter, while the A*ocs* element in A*ocs*A*mas*P*mas* is slightly shorter and corresponds to the −333/−116 domain of the same promoter. Moreover, the p*MAS* element in [u*OCS*]p*MAS* was also larger in comparison to the A*ocs*A*mas*P*mas*, with the former corresponding to the −1/−503 domain and the latter to the −1/−382 domain (numbers relative to the translation start of mannopine synthase 2′ transcript). We also constructed a shortened version of [u*OCS*]p*MAS*, called [u*OCS*s]p*MAS*s in which the u*OCS* and p*MAS* domains corresponded exactly to those in A*ocs*A*mas*P*mas* (construct pFAJ3386). The median of the lnGUS activity was significantly lower when compared to the results obtained from pFAJ3386 (8.78) to those of pFAJ3184

Corteva Exhibit 2067 - Page 6 of 11
Inari v. Corteva - PGR2023-00022

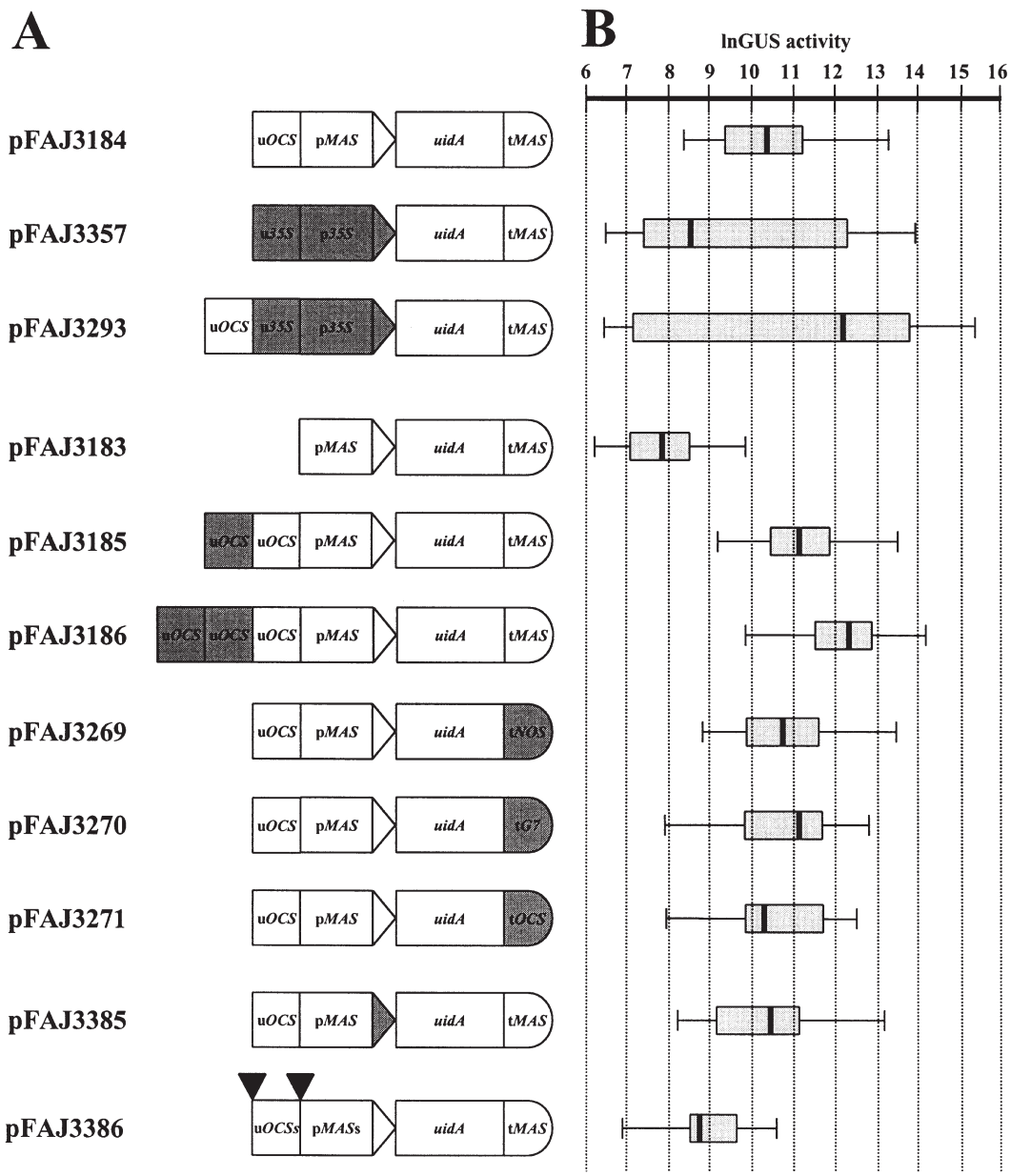

Fig. 4. (A) Schematic representation of the GUS reporter gene cassette in transformation vectors pFAJ3184, pFAJ3357, pFAJ3293, pFAJ3183, pFAJ3185, pFAJ3186, pFAJ3269, pFAJ3270, pFAJ3271, pFAJ3385, and pFAJ3386 as in Fig. 1. Abbreviations are as in Fig. 1, with addition of the following: t*G7*: T-DNA gene 7 terminator. Compared with the pFAJ3184 GUS reporter gene cassette, the differences in regulatory elements are highlighted; deletions of regulatory sequences are indicated by a black triangle. (B) Box plots of the natural logarithm of the GUS activity (lnGUS) of first generation transgenic *A. thaliana* plants transformed with the vectors pFAJ3184, pFAJ3357, pFAJ3293, pFAJ3183, pFAJ3185, pFAJ3186, pFAJ3269, pFAJ3270 pFAJ3271, pFAJ3385 and pFAJ3386 as in Fig. 1.

(10.35) whereas the variance of the lnGUS activity remained statistically invariant.

3.3. The effect of the mannopine synthase promoter on the reporter gene expression in different parts of the plant

When compared to [u*35S*]p*35S*, the use of [u*OCS*]p-*MAS* driving the GUS reporter gene ensures a constitutive level of expression with little inter-transformant variability in leaves of transgenic *A. thaliana* plants. To assess whether this expression characteristic also applies to other tissues as well, GUS expression levels were measured in leaves, stems and roots of first generation *A. thaliana* plants transformed with pFAJ3184 ([u*OCS*]p*MAS*) and pFAJ3163 ([p*35S*]p*35S*) (Table 3). In plants transformed with pFAJ3163, the median and variance of the lnGUS activity in the different tissues were comparable. In plants transformed with pFAJ3184, the median lnGUS activity was significantly higher in roots than in leaves and stems. The variance of

Corteva Exhibit 2067 - Page 7 of 11
Inari v. Corteva - PGR2023-00022

Table 2
Statistical analysis of expression levels in T0 generation populations of *A. thaliana* plants transformed with the constructs shown in Fig. 4A

Construct	No[a]	GUS activity[b]		lnGUS activity[c]			
		Median	Mean	Median	P^d	Variance	P^e
pFAJ3184	33	31 265	90 840	10.35	a	2.09	a
pFAJ3357	57	5377	150 590	8.59	b	6.15	b
pFAJ3293	35	187 212	760 690	12.14	c	12.09	b
pFAJ3183	24	2440	4100	7.80	b	1.06	a
pFAJ3185	33	64 860	137 490	11.08	a	1.41	a
pFAJ3186	35	215 345	283 450	12.28	c	1.12	a
pFAJ3269	18	44 355	121 260	10.70	a	1.91	a
pFAJ3270	30	66 171	81 460	11.10	a	1.39	a
pFAJ3271	24	29 436	73 430	10.29	a	1.71	a
pFAJ3385	33	32 450	68 110	10.39	a	1.75	a
pFAJ3386	15	6502	12 530	8.78	b	1.08	a

[a] Number of plants analyzed in the T0 generation of transformants.
[b] GUS activity is expressed in pmoles 4-methylumbelliferone per min per mg soluble protein.
[c] Natural logarithm of GUS activity.
[d] Comparisons of median at a significance level 0.05. 'a' indicates that the median is not significantly different from pFAJ3184. 'b' indicates that the median is significantly lower than pFAJ3184. 'c' indicates that the median is significant higher than pFAJ3184.
[e] Comparisons of variance at a significance level 0.05. 'a' indicates that the variance is not significantly different from pFAJ3184. 'b' indicates that the variance is significant higher than pFAJ3184.

the lnGUS activity, however, was comparable and not significantly altered in the different tissues.

4. Discussion

Although there is ample literature on the effect of MARs on the expression variability of reporter genes (reviewed by Allen et al. [14]), the influence of MARs on expression variability in *A. thaliana* has not been analyzed intensely so far. It can be concluded from the experiments in this paper that the chicken lysozyme MARs have no effect on either reporter gene expression level or variability in *A. thaliana*. The same MARs positioned in the same relative orientation had previously been shown to drastically reduce expression variability in transgenic tobacco plants [8,22]. This discrepancy might imply that the effect of MARs on transgene expression can be different from one species to another and is therefore unpredictable if not tested in a particular species. The reason for the species-dependent differences of MAR effects is not known but it must be assumed that nuclear factors interacting with MARs differ among different plant species. Also, the discrepancy of the results could be due to homology-dependent gene silencing, which might overrule the protecting effect of MARs when multiple copies of the transgene are integrated [42]. However, in the present study the T-DNA copy numbers or integration patterns were not determined, as this was considered irrelevant in respect

Table 3
Statistical analysis of expression levels in different plant parts of T0 generation populations of *A. thaliana* plants transformed with constructs shown in Fig. 1A

Construct	Tissue	No[a]	GUS activity[b]		lnGUS activity[c]			
			Median	Mean	Median	P^d	Variance	P^e
pFAJ3163	leaf	15	2253	322 630	7.72	a	8.56	a
	root	14	2416	344 360	7.79	a	8.61	a
	stem	14	7781	490 680	8.94	a	7.4	a
pFAJ3184	leaf	14	179 870	183 350	12.10	a	1.98	a
	root	12	1 520 930	417 470	14.23	b	2.68	a
	stem	11	152 000	167 240	11.93	a	1.72	a

[a] Number of plants analyzed in the T0 generation of transformants.
[b] GUS activity is expressed in pmoles 4-methylumbelliferone per min per mg soluble protein.
[c] Natural logarithm of GUS activity.
[d] Comparisons of median per construct at a significance level 0.05. Identical letters indicate that the medians are not significantly different.
[e] Comparisons of variance per construct at a significance level 0.05. Identical letters indicate that the variances are not significantly different.

Corteva Exhibit 2067 - Page 8 of 11
Inari v. Corteva - PGR2023-00022

to our goal to generate populations of transgenic plants providing minimal inter-individual variation of transgene expression.

Furthermore, it can be concluded from our experiments that a low level of inter-transformant variability can be achieved by using a promoter derived from the *A. tumefaciens* mannopine synthase gene. By itself, the mannopine synthase promoter (p*MAS*) results in a median level of GUS expression in *A. thaliana* leaves that is only slightly lower when compared to the enhanced CaMV 35S promoter ([u*35S*]p*35S*), the most frequently used promoter for constitutive transgene expression in plants. However, the variance of lnGUS expression driven by the p*MAS* is at least 8-fold lower when compared to [u*35S*]p*35S*. Even better results are obtained when an upstream activating sequence from the octopine synthase gene ([u*OCS*]) is placed upstream of p*MAS*. With this chimeric promoter, [u*OCS*]p*MAS*, the GUS expression level median is at nearly 6-fold higher and the variance of lnGUS activity 3-fold lower than with [u*35S*]p*35S*. This [u*OCS*] element, however, does not appear to be responsible for the low inter-transformant variability of expression, since addition of [u*OCS*] upstream of [u*35S*]p*35S* even further increased expression variability. The [u*OCS*] element clearly acts as an expression enhancer. Addition of one, two or three copies of [u*OCS*] upstream of p*MAS* increased the median GUS expression levels 13-, 26- and 88-fold, respectively, whereas addition of the [u*OCS*] element upstream of the [u*35S*]p*35S* boosted median expression level 60-fold.

Ni et al. [41] have previously shown that A*ocs*A*mas*P*mas* and (A*ocs*)3A*mas*P*mas*, promoters that are highly similar to the [u*OCS*]p*MAS* and [u*OCS*]3p*MAS* promoters used in this study, provide an average expression level in leaves of transgenic tobacco plants that is 46- and 75-fold higher than that obtained with the CaMV 35S promoter. However, variability of the reporter gene expression driven by these two types of promoters has not been compared. It should be noted however, that promoters A*ocs*A*mas*P*mas* and [u*OCS*]p*MAS* are not exactly similar, as [u*OCS*]p*MAS* contains a 6 bp longer portion of the octopine gene and a 119 bp longer fragment of the mannopine gene relative to the promoter fragments used in A*ocs*A*mas*P*mas*. In a direct comparison of promoter [u*OCS*s]p*MAS*s, in which these extra sequences were deleted, the [u*OCS*]p*MAS* clearly performed better than the A*ocs*A*mas*P*mas*, as median expression levels were 5-fold higher.

The reason why the p*MAS*-based promoters result in lower expression variability is currently unknown. When comparing the lnGUS expression in populations transformed with p*MAS*-based and p*35S*-based promoters, one striking difference is that p*MAS*-driven gene expression is characterized by a normal distribution with a single class of moderately high expression levels, whereas p*35S*-driven expression results in a bimodal distribution with about 2/3 of the population in a low expression state and the remaining 1/3 in a high expression state (Fig. 3). Other research groups also observed an analogous bimodal expression pattern for p*35S* [7,43]. This could indicate that silencing phenomena occur less frequent in plants transformed with p*MAS* versus p*35S*. It has been stated that viral and bacterial DNA sequences can trigger silencing processes more frequently than plant-derived sequences [9]. Hence, the p*35S* promoter may be more prone to silencing compared to plant-derived promoters simply by virtue of its viral origin. However, this view is contradicted by the observations that the potato *Lhca3.St1* promoter caused higher expression variability in first generation transgenic tobacco plants than the enhanced p35S promoter [44] and that the *A. thaliana* UBQ1 promoter likewise caused higher expression variability than the p35S promoter in first generation *A. thaliana* plants [45]. Hence, the low expression variability conferred by the p*MAS*-derived promoters appears to be more exceptional than the high expression variability observed with the CaMV 35S promoter. One possible explanation for the low expression variability observed with p*MAS*-derived promoters, yet completely speculative at present, might be that its key *cis*-acting elements are not easily methylated. Methylation of promoter sequences has been associated in many cases with silencing effects [46] and is therefore believed to contribute to inter-transformant expression variability.

The influence of other regulatory elements on the expression variability of a reporter gene in *A. thaliana* has so far been poorly investigated. From our experiments, we can conclude that the expression characteristics of the reporter gene driven by [u*OCS*]p*MAS* were not influenced by any of four types of terminators (t*MAS*, t*NOS*, t*G7*, t*OCS*) nor by any of two different 5′ untranslated regions (5′ untranslated region of mannopine synthase or 5′ untranslated region of tobacco mosaic virus). Hence, the role of promoter elements in p*MAS* seems to be predominant.

The ability of the [u*OCS*]p*MAS* promoter to confer weakly variable, constitutive expression in *A. thaliana* was not only observed in leaves but also in stems and roots, making it the promoter of choice for a wide range of applications. This promoter can be used to study the effect of constitutive over-expression of particular genes in *A. thaliana* plants. Nearly every individual in a population of first generation transgenic plants carrying a transgene driven by [u*OCS*]p*MAS* will show a relatively high level of expression, and hence only one or a few transgenic plants are required to study the phenotypic effect of over expression. This markedly facilitates applications in which high numbers of cDNA's have to be tested in constitutive over expression mode, for instance for functional analysis of the

Corteva Exhibit 2067 - Page 9 of 11
Inari v. Corteva - PGR2023-00022

thousands of *A. thaliana* genes to which no function has yet been assigned.

Acknowledgements

This work was supported partly by a grant from the Flanders Interuniversity Institute for Biotechnology (PRJ3) and partly by a grant from the Fonds voor Wetenschappelijk Instituut Vlaanderen (G.0118.01). K.M.J. Butaye is indebted to the Instituut voor de Aanmoediging van Innovatie door Wetenschap en Technologie in Vlaanderen for a predoctoral fellowship. We wish to thank Dr W.H. Strätling for pUC-B-1-X1.

References

[1] S.L.A. Hobbs, P. Kpodar, C.M.O. DeLong, The effect of T-DNA copy number and methylation on reporter gene expression in tobacco tranformants, Plant Mol. Biol. 15 (1990) 851–864.
[2] L. Jones, A.J. Hamilton, O. Voinnet, C.L. Thomas, A.J. Maule, D.C. Baulcombe, RNA–DNA interactions and DNA methylation in post-transcriptional gene silencing, Plant Cell 11 (1999) 2291–2302.
[3] C. Peach, J. Velten, Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters, Plant Mol. Biol. 17 (1991) 49–60.
[4] B. Tinland, The integration of T-DNA into plant genomes, Trends Plant Sci. 1 (1996) 178–184.
[5] A.J.M. Matzke, M.A. Matzke, Position effects and epigenetic silencing of plant genes, Curr. Opin. Plant Biol. 1 (1998) 142–148.
[6] A. Depicker, M. Van Montagu, Post-transcriptional gene silencing in plants, Curr. Opin. Cell Biol. 7 (1997) 373–382.
[7] S.L.A. Hobbs, T.D. Warkentin, C.M.O. DeLong, Transgene copy number can be positively or negatively associated with transgene expression, Plant Mol. Biol. 21 (1993) 17–26.
[8] L. Mlynárová, A. Loonen, J. Heldens, R.C. Jansen, P. Keizer, W.J. Stiekema, J.P. Nap, Reduced position effect in mature transgenic plants conferred by the chicken lysozyme matrix-associated region, Plant Cell 6 (1994) 417–426.
[9] J.M. Kooter, M.A. Matzke, P. Meyer, Listening to the silent genes: transgene silencing, gene regulation and pathogen control, Trends Plant Sci. 4 (1999) 340–347.
[10] C. Ma, A. Mitra, Intrinsic direct repeats generate consistent post-transcriptional gene silencing in tobacco, Plant J. 31 (2002) 37–39.
[11] P. Meyer, H. Saedler, Homology-dependent gene silencing in plants, Annu. Rev. Plant Physiol. Plant Mol. Biol. 47 (1996) 23–48.
[12] S. Spiker, W.F. Thompson, Nuclear matrix attachment regions and transgene expression in plants, Plant Physiol. 110 (1996) 15–21.
[13] R. Holmes-Davis, L. Comai, Nuclear matrix attachment regions and plant gene expression, Trends Plant Sci. 3 (1998) 91–97.
[14] G.C. Allen, S. Spiker, W.F. Thompson, Use of matrix attachment regions (MARs) to minimize transgene silencing, Plant Mol. Biol. 43 (2000) 361–376.
[15] U.K. Laemmli, E. Käs, L. Poljak, Y. Adachi, Scaffold-associated regions: *cis*-acting determinants of chromatin structural loops and functional domains, Curr. Opin. Genet. Dev. 2 (1992) 275–285.
[16] J.-W. Liu, L.M. Tabe, The influence of two plant nuclear matrix attachment regions (MARs) on gene expression in transgenic plants, Plant Cell Physiol. 39 (1998) 115–123.
[17] K. Petersen, R. Leah, S. Knudsen, V. Cameron-Mills, Matrix attachment regions (MARs) enhance transformation frequencies and reduce variance of transgene expression in barley, Plant Mol. Biol. 49 (2002) 45–58.
[18] F. Schöffl, G. Schröder, M. Kliem, M. Rieping, An SAR sequence containing 395 bp DNA fragment mediates enhanced, gene-dosage-correlated expression of chimaeric heat shock gene in transgenic tobacco plants, Transgenic Res. 2 (1993) 93–100.
[19] B. Ülker, G.C. Allen, W.F. Thompson, S. Spiker, A.K. Weissinger, A tobacco matrix attachment region reduces the loss of transgene expression in the progeny of transgenic tobacco plants, Plant J. 18 (1999) 253–263.
[20] P. Vain, B. Worland, A. Kohli, J.W. Snape, P. Christou, G.C. Allen, W.F. Thompson, Matrix attachment regions increase transgene expression levels and stability in transgenic rice plants and their progeny, Plant J. 18 (1999) 233–242.
[21] A.H.M. Van der Geest, G.E. Hall, Jr., S. Spiker, T.C. Hall, The β-phaseolin gene is flanked by matrix attachment regions, Plant J. 6 (1994) 413–423.
[22] L. Mlynárová, R.C. Jansen, W. Conner, W.J. Stiekema, J.-P. Nap, The MAR-mediated reduction in position effect can be uncoupled from copy number-dependent expression in transgenic plants, Plant Cell 7 (1995) 599–609.
[23] E.M. Meyerowitz, Arabidopsis a useful weed, Cell 27 (1989) 263–269.
[24] I.J.W. Goderis, M.F.C. De Bolle, I.E.J.A. François, P.F.J. Wouters, W.F. Broekaert, B.P.A. Cammue, A set of modular plant transformation vectors allowing flexible insertion of up to six expression units, Plant Mol. Biol. 50 (2002) 17–27.
[25] L. Phi-Van, W.H. Strätling, The matrix attachment regions of the chicken lysozyme gene co-map with the boundaries of the chromatin domain, EMBO J. 7 (1988) 655–664.
[26] R.S.S. Datla, F. Bekkaoui, J.K. Hammerlindl, G. Pilate, I.D. Dunstan, W.L. Crosby, Improved high-level constitutive foreign gene expression using an AMV RNA4 untranslated leader sequence, Plant Sci. 94 (1993) 139–149.
[27] R.C. Gardner, A.J. Howarth, P. Hahn, M. Brown-Luedi, R.J. Shepperd, J. Messing, The complete nucleotide sequence of an infectious clone of *cauliflower mosaic virus* by M13mp7 shotgun sequencing, Nucl. Acids Res. 9 (1981) 2871–2888.
[28] R.F. Barker, K.B. Idler, D.V. Thompson, J.D. Kemp, Nucleotide sequence of the T-DNA region from *Agrobacterium tumefaciens* octopine Ti plasmid pTi15955, Plant Mol. Biol. 2 (1983) 335–350.
[29] D.R. Gallie, V. Walbot, Identification of the motifs within the tobacco mosaic virus 5′-leader responsible for enhancing translation, Nucl. Acids Res. 17 (1992) 4631–4638.
[30] A. Depicker, S. Stachel, P. Dhaese, P. Zambryski, H.M. Goodman, Nopaline synthase: transcript mapping and DNA sequence, J. Mol. Appl. Gen. 1 (1982) 561–573.
[31] P. Zambryski, H. Joos, C. Genetello, J. Leemans, M. Van Montagu, J. Schell, Ti plasmid vectors for the introduction of DNA into plant cells without alteration of their normal regeneration capacity, EMBO J. 2 (1983) 2143–2150.
[32] N. Bechthold, J. Ellis, G. Pelletier, In *planta Agrobacterium*-mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants, C.R. Acad. Sci. Paris 316 (1993) 1194–1199.
[33] K. Eggermont, I.J. Goderis, W.F. Broekaert, High-throughput RNA extraction from plant samples based on homogenization by reciprocal shaking in the presence of a mixture of sand and glass beads, Plant Mol. Biol. Rep. 14 (1996) 273–279.
[34] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, second ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989.

Corteva Exhibit 2067 - Page 10 of 11
Inari v. Corteva - PGR2023-00022

[35] R.A. Jefferson, Assaying chimeric genes in plants: the GUS gene fusion system, Plant Mol. Biol. Rep. 5 (1987) 387–405.
[36] M.M. Bradford, A rapid and sensitive method for the quantification of microgram quantities utilizing the principle of protein-dye binding, Anal. Biochem. 72 (1976) 248–254.
[37] J.-P. Nap, P. Keizer, R. Jansen, First-generation transgenic plants and statistics, Plant Mol. Biol. Rep. 11 (1993) 156–164.
[38] B. Efron, R.J. Tibshirani, An Introduction to the Bootstrap Monographs on Statistics and Applied Probability, vol. 75, Chapman and Hall, New York, 1993.
[39] D.D. Boos, C. Brownie, Bootstrap methods for testing homogeneity of variances, Technometrics 31 (1989) 69–82.
[40] P.H. Westfall, S.S. Young, Resampling-Based Multiple Testing, Wiley, New York, 1993.
[41] M. Ni, D. Cui, J. Einstein, S. Narasimhulu, C.E. Vergara, S.B. Gelvin, Strength and tissue specificity of chimeric promotors derived from the octopine and mannopine synthase genes, Plant J. 7 (1995) 661–676.
[42] H. Vaucheret, T. Elmayan, D. Thierry, A. van der Geest, T. Hall, A.J. Corner, L. Mlynárová, J.P. Nap, Flanking matrix attachment regions (MARs) from chicken, bean, yeast or tobacco do not prevent homology-dependent *trans*-silencing in transgenic tobacco plants, Mol. Gen. Genet. 259 (1998) 388–392.
[43] T. Elmayan, S. Balzergue, F. Béon, V. Bourdon, J. Daubremet, Y. Guénet, P. Mourrain, J.-C. Palauqui, S. Vernhettes, T. Vialle, K. Wostrikoff, H. Vaucheret, Arabidopsis mutants impaired in cosuppression, Plant Cell 10 (1998) 1747–1757.
[44] L. Mlynárová, L.C.P. Keizer, W.J. Stiekema, J.-P. Nap, Approaching the lower limits of transgene variability, Plant Cell 8 (1996) 1589–1599.
[45] S. Holtorf, K. Apel, H. Bohlman, Comparison of different constitutive and inducible promoters for the overexpression of transgenes in *Arabidopsis thaliana*, Plant Mol. Biol. 29 (1995) 637–646.
[46] J.D.G. Jones, P. Dunsmuir, J. Bedbrook, High level expression of introduced chimaeric genes in regenerated transformed plants, EMBO J. 4 (1985) 2411–2418.

Corteva Exhibit 2067 - Page 11 of 11
Inari v. Corteva - PGR2023-00022