The Plant Cell, Vol. 7, 599–609, May 1995 © 1995 American Society of Plant Physiologists

The MAR-Mediated Reduction in Position Effect Can Be Uncoupled from Copy Number-Dependent Expression in Transgenic Plants

Ľudmila Mlynárová,[a] Ritsert C. Jansen,[b] Anthony J. Conner,[a,1] Willem J. Stiekema,[a] and Jan-Peter Nap [a,2]

[a] Department of Molecular Biology, Centre for Plant Breeding and Reproduction Research (CPRO-DLO), P.O. Box 16, NL-6700 AA Wageningen, The Netherlands

[b] Department of Population Biology, Centre for Plant Breeding and Reproduction Research (CPRO-DLO), P.O. Box 16, NL-6700 AA Wageningen, The Netherlands

To study the role of matrix-associated regions (MARs) in establishing independent chromatin domains in plants, two transgenes were cloned between chicken lysozyme A elements. These transgenes were the neomycin phosphotransferase (NPTII) gene under control of the nopaline synthase (nos) promoter and the β-glucuronidase (GUS) gene controlled by the double cauliflower mosaic virus (dCaMV) 35S promoter. The A elements are supposed to establish an artificial chromatin domain upon integration into the plant DNA, resulting in an independent unit of transcriptional regulation. Such a domain is thought to be characterized by a correlated and position-independent, hence copy number-dependent, expression of the genes within the domain. The presence of MARs resulted in a higher relative transformation efficiency, demonstrating MAR influence on NPTII gene expression. However, variation in NPTII gene expression was not significantly reduced. The selection bias for NPTII gene expression during transformation could not fully account for the lack of reduction in variation of NPTII gene expression. Topological interactions between the promoter and A element may interfere with the A element as a domain boundary. In contrast, the GUS gene on the same putative chromatin domain showed a highly significant reduction in variation of gene expression, as expected from previous results. Surprisingly, no copy number-dependent GUS gene expression was observed: all plants showed approximately the same GUS activity. We concluded, therefore, that dCaMV 35S-GUS gene expression in mature tobacco plants is regulated by some form of dosage compensation.

INTRODUCTION

Genetic transformation of plants generally results in individual transformants with highly variable levels of transgene expression (Peach and Velten, 1991; Nap et al., 1993a). One extreme in this wide range of gene expression levels is the fully silenced transgene; examples have been better documented in recent years (Kooter and Mol, 1993; Finnegan and McElroy, 1994; Flavell, 1994). The variation between transformants hampers the analysis of plant regulatory sequences and the exploitation of plant genes by the transgenic approach. More insight into the cause and characteristics of this variation is likely to be important for our understanding of the regulation of (trans)-gene expression in a plant genome. The between transformant variation is usually attributed to position effects due to differences in the integration site of the newly introduced DNA (Dean et al., 1988; Peach and Velten, 1991) and/or to differences in T-DNA copy number or T-DNA organization (Breyne et al., 1992b; Assaad et al., 1993; Hobbs et al., 1993).

The presence of matrix-associated regions (MARs) around reporter genes has been shown to reduce the variability of transgene expression in animal cells (Stief et al., 1989) as well as in plant calli (Breyne et al., 1992a). We have previously demonstrated that one of the better characterized animal MARs, the chicken lysozyme A element, has specific affinity for tobacco nuclear matrices and lacks general enhancer activity in plant cells (Mlynárová et al., 1994). Cloning the A element around the β-glucuronidase (GUS) gene (*uidA*), under control of a promoter from the potato light-harvesting complex-related *Lhca3.St.1* gene, showed slight reduction in the variability of GUS gene expression in mature transgenic tobacco plants (Mlynárová et al., 1994). Similar reductions have recently been obtained with bean phaseolin MAR sequences in transgenic tobacco (van der Geest et al., 1994). In contrast, when the A elements were placed just inside the T-DNA borders, flanking the GUS reporter gene as well as the

[1] Current address: New Zealand Institute for Crop & Food Research Ltd., Private Bag 4704, Christchurch, New Zealand.

[2] To whom correspondence should be addressed.

neomycin phosphotransferase II (NPTII) selection gene (*aphA2*), we have observed a much stronger reduction in the variability of GUS gene expression (Mlynárová et al., 1994). In addition, the reduced variability revealed a strong correlation between copy number and level of GUS gene expression for plants carrying up to two copies of the T-DNA. These data support the DNA loop model for the regulation of plant (trans)gene expression.

The DNA loop model for eukaryotic gene expression is based upon in vitro experiments with animal interphase nuclei or mitotic chromosomes. After selective extraction of histones, independent loops of DNA could be observed. MAR elements are thought to form the boundaries of these DNA loops in the interphase nucleus. This way, they either insulate genes within the loop from the influences of surrounding chromatin (Laemmli et al., 1992) or prevent mislocalization of genes into heterochromatin (Dorer and Henikoff, 1994). This shielding of the transgene(s) is predicted to yield position-independent expression of transgenes, resulting in a copy number–dependent expression of transgenes (Stief et al., 1989; Schlake et al., 1994). The DNA loop hypothesis furthermore predicts that flanking transgenes with MAR elements will create a more independent, coordinatedly regulated transcriptional unit of gene regulation.

Here, we address these predictions of the DNA loop hypothesis for transgene expression in mature transgenic plants by analyzing gene expression of two genes placed on an artificial MAR-flanked chromatin domain. Our results showed that a significant reduction in variation of gene expression was conferred upon the GUS gene driven by the double cauliflower

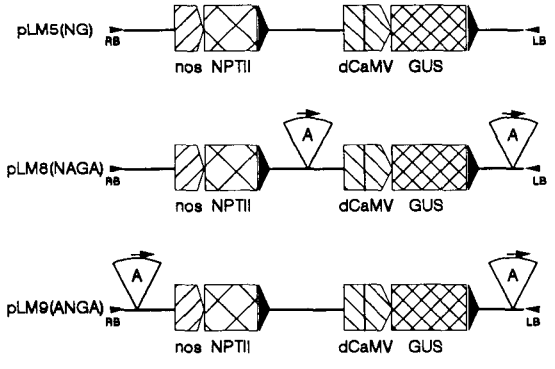

Figure 1. Overview of the T-DNA Region of Vectors Used in This Study.

For all three plant transformation vectors, the relative positions and orientations of the GUS gene, the dCaMV 35S promoter, the nos promoter, the NPTII gene, and the two nos terminators (black triangles) are shown with respect to the left border (LB) and right border (RB) of the T-DNA. The chicken lysozyme A element is indicated as an open triangle labeled A. The arrows on top of the open triangles show the orientation of the A element. The T-DNA regions are not drawn to scale. Detailed cloning steps are described in Methods. To facilitate identification, the T-DNA configuration is added to the name of the plasmid: N, NPTII gene; G, GUS gene; and A, A element, yielding pLM5(NG), pLM8(NAGA), and pLM9(ANGA).

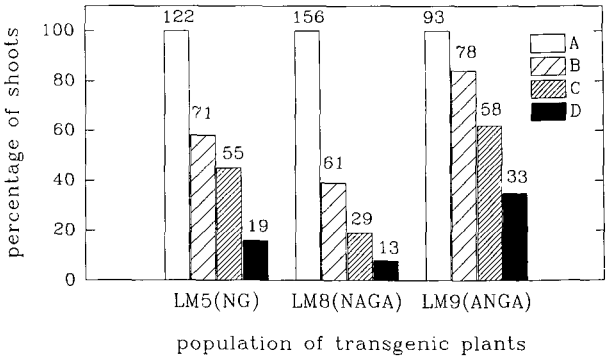

Figure 2. Relative Transformation Efficiencies.

For each transformation vector, the total number of shoots tested for rooting on 50 μg/mL kanamycin (bars A) is given as 100%. The percentage of shoots relative to this number is plotted. The open bars (A) represent the total number of shoots excised from the leaf discs after transformation with Agrobacterium and assayed for rooting. The hatched bars (B) represent the percentage of shoots rooting in the presence of 50 μg/mL kanamycin, the hatched bars (C) represent the percentage of plants with at least one intact GUS gene, and the black bars (D) represent the percentage of plants with one intact copy of the T-DNA as determined by DNA gel blot analyses. The numbers on top of the bars indicate the actual number of shoots to which the percentages refer. The values for bars A do not reflect the total number of shoots formed on the leaf discs. The presence of both border fragments was not determined for all plants (see Tables 2 and 3); therefore, the percentage and number of shoots given for bars D are corrected for the relative number of shoots from which the DNA was examined for both border fragments by DNA gel blot analysis. As a result, the numbers are not identical to the numbers given in Table 2. The populations of plants were named according to the T-DNA construct that they contained.

mosaic virus (dCaMV) 35S promoter, but this reduction was not accompanied by a copy number dependence of gene expression. Despite the presence of the chromatin boundary elements, the two genes within the presumed loop showed no correlated regulation of expression.

RESULTS

Improved Relative Transformation Efficiency

The T-DNA of the plant transformation vectors used in this study are depicted in Figure 1. In contrast to the transformation vectors described previously (Mlynárová et al., 1994), the orientations of the promoter–gene cassettes and the A elements in the three vectors were now identical with respect to the T-DNA borders. For each of these three plant transformation vectors, more than 60 independent tobacco transformants were regenerated. Shoots that rooted in the presence of 50 μg/mL kanamycin were considered transgenic. Because the NPTII selection gene is located adjacent to the right T-DNA

border sequence, it is possible that not all transgenic plants carry the GUS reporter gene. Therefore, all plants shown by DNA gel blot analyses not to contain an intact copy of the GUS gene were discarded from subsequent analyses. Overall, tobacco shoot regeneration was too frequent to permit accurate determination of transformation frequencies relative to the number of leaf discs cocultivated. Therefore, alternative measures of transformation efficiencies were used.

In Figure 2, the proportions of shoots rooting in the presence of kanamycin, of plants carrying the GUS gene, and of plants carrying one complete T-DNA are presented (see following discussion). A binomial test (Sokal and Rohlf, 1981) showed that pLM9(ANGA) transformed significantly better than did the other two T-DNA constructs, taking either rooted ($P < 0.001$) or GUS gene–containing ($P < 0.05$) plantlets as an indication of transformation efficiency. The total number of shoots that regenerated from leaf discs transformed with pLM9(ANGA) was less than the number regenerating from discs transformed with pLM8(NAGA) or pLM5(NG) (data not shown). Despite this result, pLM9(ANGA) yielded the highest number (and percentage) of transgenic plants with one intact copy of the T-DNA (Figure 2; $P < 0.001$). We also observed this result with a similar construct having the light-harvesting complex–related promoter (Mlynárová et al., 1994). Apparently, the transformation with pLM9(ANGA) results in a higher frequency of single, intact T-DNA integrations.

Reduced Variation in GUS Gene but Not in NPTII Gene Expression

GUS activity was determined in leaves of greenhouse-grown plants. Three plants from the LM5(NG) population and one plant from the LM8(NAGA) population contained the GUS gene but showed no GUS gene expression. Although the results with these individual plants are shown in Figure 3, they were omitted from the statistical evaluations (Table 1). Compared with the LM5(NG) control population, the LM9(ANGA) population showed a significant, sevenfold reduction in the variation of GUS gene expression levels (Table 1). Similarly, the LM8(NAGA) population showed a twofold reduction. Although mean GUS activity was significantly higher in the LM8(NAGA) and LM9(ANGA) populations (Table 1), maximum activity was similar in all three populations (Figure 3).

In most plants expressing the GUS gene, the amount of NPTII protein was determined with an ELISA-based assay. Amounts of NPTII protein were determined in leaves of in vitro–grown plantlets. In greenhouse-grown plants, the background was too high when this assay was used. In four plants from the LM5(NG) population, one plant from the LM8(NAGA) population, and three plants from the LM9(ANGA) population, the NPTII gene was present without resulting in detectable amounts of NPTII. The results with these individual plants are shown in Figure 4 but were omitted from the statistical analyses (Table

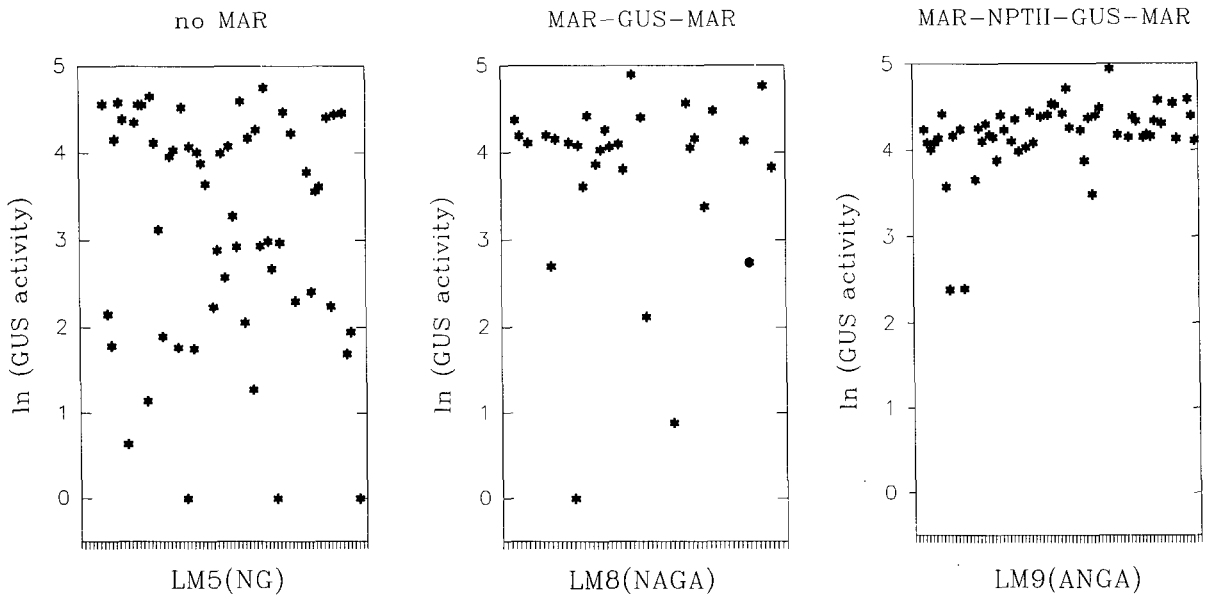

Figure 3. Decreased Variability of GUS Activity.

The natural logarithm (ln) of GUS activity of each individual transformant known to contain the GUS gene is plotted for each plant population, named according to the T-DNA construct contained. The value plotted is the mean ($n = 4$) of the natural logarithms of the activities in leaves of greenhouse-grown plants. Each point on the *x*-axis represents an individual transformant. For each population, the T-DNA configuration is given above the graphs. The statistics describing the three populations are given in Table 1.

Table 1. Overall Statistics of the Three Populations

Gene Population[a]	No.[b]	Mean[c]	SE[d]	P[e]	Var.[f]	P[g]	Fold[h]
GUS							
LM5(NG)	54	3.34	0.15	—[i]	1.26	—	—
LM8(NAGA)	29	3.88	0.15	**[j]	0.69	*[j]	2
LM9(ANGA)	55	4.17	0.06	***[j]	0.19	***	7
NPTII							
LM5(NG)	41	7.58	0.14	—	0.78	—	—
LM8(NAGA)	20	7.17	0.12	*	0.27	**	3
LM9(ANGA)	38	7.73	0.14	NS[k]	0.73	NS	NA[l]

[a] Populations are named according to the T-DNA construct the plants contain.
[b] Number of plants in the population.
[c] Means are based on the natural logarithms of enzyme activities (GUS gene) or amounts of protein (NPTII gene).
[d] Standard error of the mean.
[e] Probability according to the *t* test for location in case of unequal variances (Sokal and Rohlf, 1981) compared with the corresponding control population LM5(NG).
[f] Variance.
[g] Probability according to the *F* test for homogeneity of variances (Sokal and Rohlf, 1981) compared with the corresponding control population LM5(NG).
[h] Fold reduction of variance with respect to the corresponding control population LM5(NG).
[i] (—), control population.
[j] Indication of probability value; ***, significant at $P < 0.001$; **, significant at $P < 0.01$; *, significant at $P < 0.05$.
[k] Not significant (at $P = 0.05$).
[l] Not applicable.

1). These analyses showed no significant reduction in the variation of NPTII gene expression in the LM9(ANGA) population when compared with the LM5(NG) population. The variation in the amounts of NPTII was reduced threefold in the LM8(NAGA) population, but this result was based on considerably fewer plants. The reduction in variation was not apparent in subpopulations having the same copy number (see following discussion). Restricting the analyses to plants carrying only intact copies of the T-DNA, as determined by equal numbers of left and right border fragments, did not affect the conclusions drawn (data not shown).

A double logarithmic plot of GUS activity versus the amount of NPTII illustrated that there was no relationship between the expression levels of the two transgenes in any of the three populations of transgenic plants analyzed (Figure 5). None of the correlation coefficients were statistically significant (at $P = 0.05$). Similar analyses with GUS activity data from in vitro–grown plantlets showed identical results (data not shown).

Relationship between Expression Levels and Copy Numbers

To establish the relationship between the number of integrated transgenes and the level of GUS gene expression, copy numbers were determined for all plants found by DNA gel blot analysis to contain the GUS gene. The number of T-DNA inserts was determined by counting the number of fragments on DNA gel blots representing junctions between plant DNA and T-DNA. This analysis was performed for both left and right

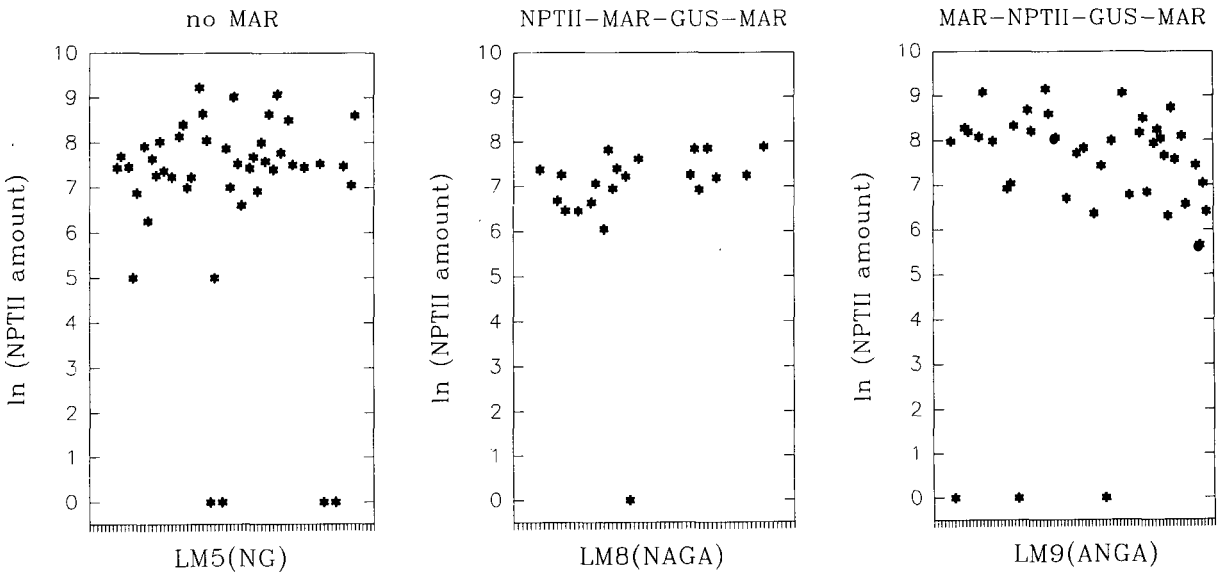

Figure 4. No Differences in Variability of the Amount of NPTII.

The natural logarithm (ln) of the amount of NPTII in each individual transformant is plotted for each plant population, named according to the T-DNA construct contained. The value plotted is the mean ($n = 4$) of the natural logarithms of the activities in leaves of in vitro–grown plants. Each point on the *x*-axis represents an individual transformant. For each population, the T-DNA configuration is given above the graphs. The statistics describing the three populations are given in Table 1.

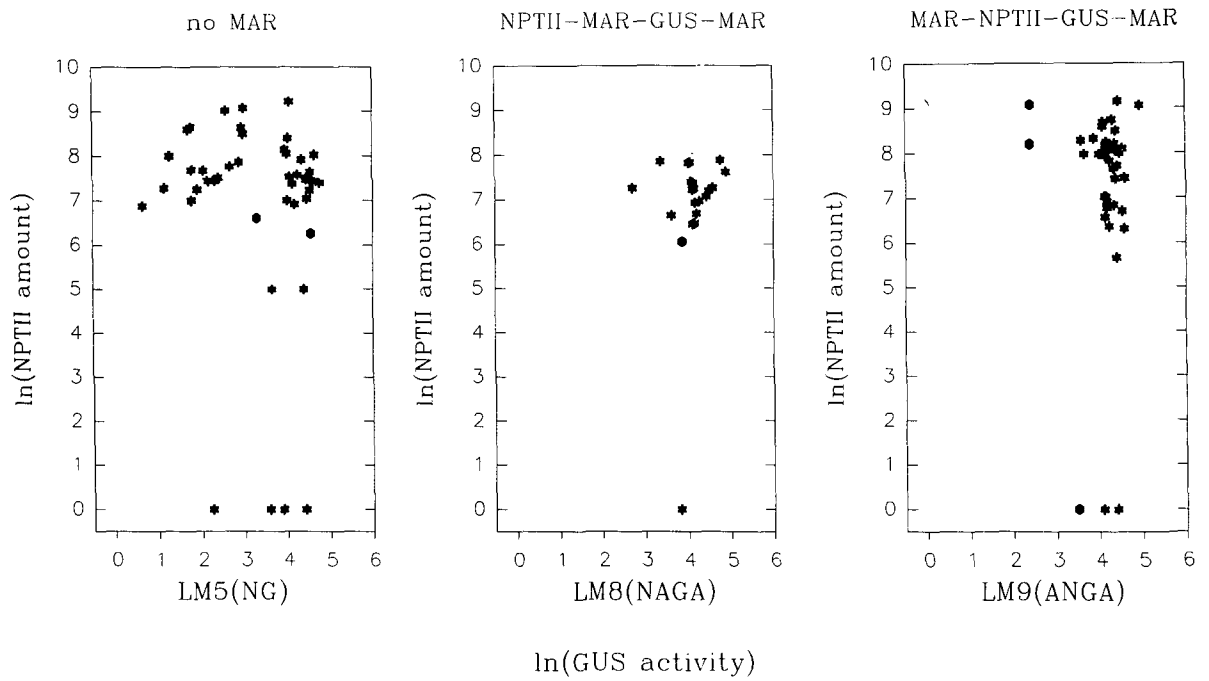

Figure 5. Relationship between GUS Activity and Amount of NPTII.

For each of the three plant populations, the natural logarithm (ln) of GUS activity is plotted against the natural logarithm of the amount of NPTII. For each population, the T-DNA configuration is given above the graphs. Correlation coeffients are −0.14 for LM5(NG), +0.37 for LM8(NAGA), and −0.21 for LM9(ANGA); all are not significant at P = 0.05.

border fragments, yielding the number of GUS and NPTII gene copies, respectively. T-DNA truncations were found in all populations, most notably in the LM8(NAGA) population (Figure 2). A large variation in GUS activity was observed in LM5(NG) plants having the same GUS gene copy numbers. Despite the large variation, there was a significant negative correlation between GUS activity and GUS gene copy number ($P < 0.001$; Figure 6). In contrast, the subpopulations of LM9(ANGA) plants showed a significant reduction in the variation of GUS gene expression (Figure 6 and Table 2). This reduction was fourfold for plants carrying one GUS gene copy and as high as 36-fold for plants carrying two GUS gene copies. However, all plants had approximately the same activity, and as a result, no relationship between GUS activity and GUS gene copy number was apparent (Figure 6). The subpopulations of the LM8(NAGA) plants showed no significant reduction in the variation of GUS activity and no correlation with copy number (Figure 6 and Table 2).

The right border fragment data were used to analyze the relationship between the amount of NPTII and NPTII gene copy number (Figure 7 and Table 3). Only in the few LM9(ANGA) plants with two copies of the NPTII gene was the slight reduction in the variation of the amount of NPTII significant. All other subpopulations were not significantly different from the corresponding LM5(NG) control population. Despite the overall high variation, the LM5(NG) copy number was positively correlated with NPTII activity ($P < 0.01$; Figure 7), but this does not hold true for the LM8(NAGA) and LM9(ANGA) populations.

DISCUSSION

The results presented here confirm the efficacy of the chicken lysozyme A element in reducing the variation of transgene expression in populations of transgenic tobacco plants. The T-DNA vector in which the chicken lysozyme A element is present at both border sequences yielded populations of primary transformants that showed highly significant and meaningful reductions of variability in transgene expression levels (Mlynárová et al., 1994; Figure 3 and Table 1). Moreover, the T-DNA vector showed a relative transformation efficiency higher than the corresponding vector without the A elements and yielded a high proportion of plants carrying one intact copy of the T-DNA (Figure 2).

Evidence is accumulating that at least part of the observed variation in transgene expression levels is due to the organization of the T-DNA in repeat structures (Flavell, 1994), notably inverted repeats (Hobbs et al., 1993). By restricting statistical analyses to those plants carrying intact copies of the T-DNA, we showed that this cause of variation is not an important explanatory variable in our populations of transgenic plants. The

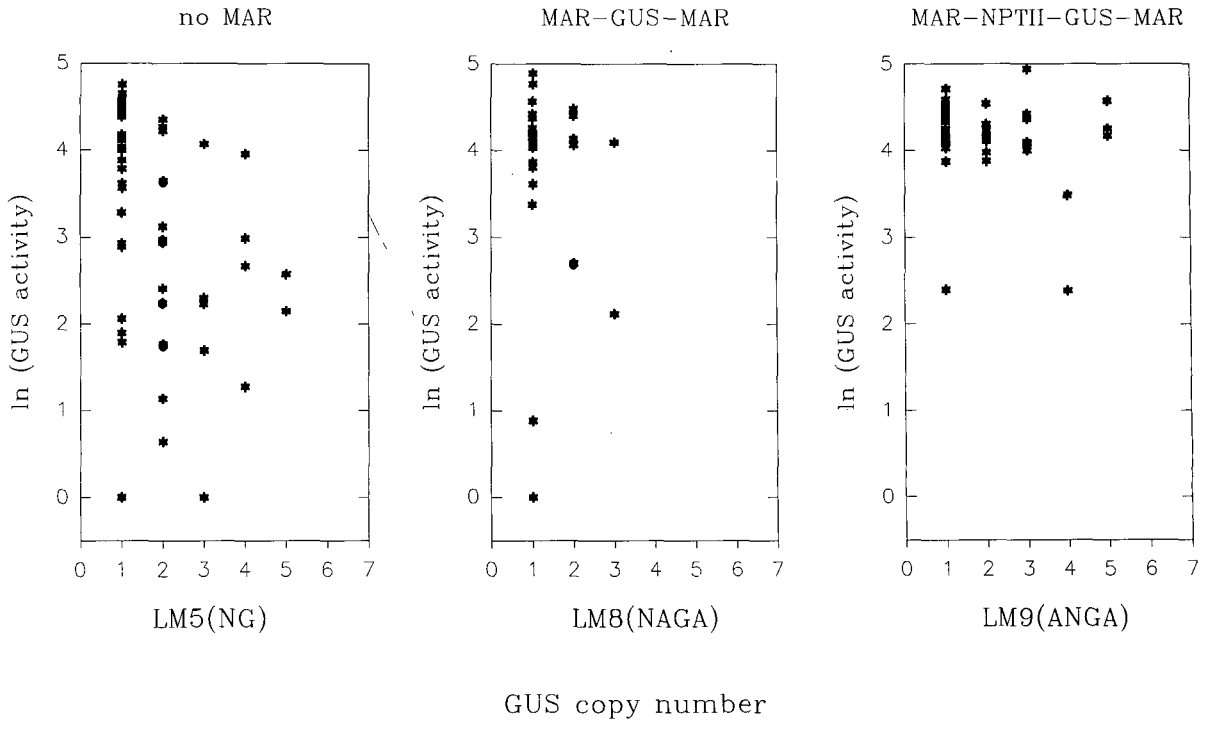

Figure 6. GUS Activity as a Function of GUS Gene Copy Number.

For each of the three plant populations, the number of GUS gene copies, as determined by DNA gel blot analysis, is plotted against the natural logarithm (ln) of GUS activity. Due to the way these data are plotted, not all points corresponding to individual plants are discernible. For each population, the T-DNA configuration is given above the graphs. The statistics describing the subpopulations of plants are given in Table 2. Linear regression analysis was performed to fit the equation ln(GUS activity) = $C + r \times$ ln(copy), with regression parameters C (a constant estimating the level of GUS activity) and r (the slope giving the correlation between GUS activity and GUS gene copy number). This analysis yielded for the LM5(NG) population a C value of 3.8 (±0.3) and an r value of −1.0 (±0.5), which was significant at $P < 0.001$; for the LM8(NAGA) population, this yielded a C value of 4.0 (±0.3) and an r value of −0.4 (±0.7), which was not significant at $P = 0.05$; and for the LM9(ANGA) population, this yielded a C value of 4.2 (±0.1) and an r value of −0.1 (±0.1), which was also not significant at $P = 0.05$.

reduction in the variation of gene expression is not promoter dependent, because similar reductions in overall variability of GUS gene expression were found with two different promoters: the light-regulated promoter from the potato gene encoding apoprotein 2 of the light-harvesting complex of photosystem I (*Lhca3.St.1*; Mlynárová et al., 1994) and a viral promoter (dCaMV 35S; this study). This indicates, as predicted by the DNA loop model, that the flanking MAR elements establish a T-DNA configuration less susceptible to influences of chromatin in the topological vicinity of the integration site. Presumably, this results from the affinity of the A element for the nuclear matrix. However, the reduced variation in gene expression was not accompanied by copy number–dependent gene expression. Also, no domain of transcriptional regulation was observed. Cloning the A element around only the GUS reporter gene proved much less efficient in reducing the variation of gene expression. These observations may challenge the DNA loop model and are addressed in the following discussion.

Reduced Position Effect without Copy Number–Dependent Gene Expression

Despite the significant reduction in the variation of GUS gene expression, virtually all plants of the LM9(ANGA) population had the same high level of GUS activity, irrespective of the number of T-DNA copies integrated (Figure 6 and Table 2). This is surprising, because position-independent gene expression is predicted to result in copy number–dependent gene expression (Stief et al., 1989; Schlake et al., 1994). In transgenic animal studies, copy number dependence of gene expression is usually taken to indicate a lack of position effects (e.g., Bonifer et al., 1990; Schedl et al., 1993). In the case of the dCaMV 35S promoter in transgenic tobacco plants, the reduction of position effects is apparently uncoupled from copy number dependence of transgene expression. The particular promoter is important because the Lhca3.St.1 promoter revealed a copy number dependence of gene expression after

the variation in gene expression was reduced (Mlynárová et al., 1994). The weaker single CaMV 35S promoter was reported to yield additive expression levels in tobacco plants (Hobbs et al., 1993) and high copy number expression in tobacco calli (Allen et al., 1993). The level of gene expression does not seem to be involved, because the Lhca3.St.1 promoter is considerably stronger than the dCaMV 35S promoter (Nap et al., 1993b).

Dosage Compensation

It could be argued that proof of position independence requires only that all plants carrying the same number of gene copies give the same level of expression (see Reitman et al., 1990). The identical expression level over a range of copy numbers then indicates that the individual copies of the transgene are expressed unequally. The lack of copy number–dependent gene expression is reminiscent of the dosage compensation observed for the alcohol dehydrogenase alleles in maize (Birchler, 1981). It may be a plant counterpart of the X-linked dosage compensation observed in various organisms; regulation of this dosage compensation was recently linked to chromatin structure (Chuang et al., 1994).

At this point, several scenarios are conceivable to explain the apparent dosage compensation. A MAR-mediated chromatin opening (Laemmli et al., 1992) may ensure optimal access for all required *trans*-acting factors, one of which is limiting. This implies that there is only a limited number of active genes (in this case, one) and that all other identical transgenes are silent. Alternatively, transgenes may interact in an as yet undetermined manner in which the number of interacting genes may be less important than the level of the final product. Silencing phenomena indicate that genes in the interphase nucleus interact and turn each other off (Kooter and Mol, 1993). Conversely, genes may interact and coordinate their expression. Possibly, the perturbation of this coordination results in *trans* inactivation. The interaction may be related to the RNA threshold level model put forward for sense suppression in plants (Smith et al., 1994). This model seeks to explain the reaction of plant cells to viral stress, but it may describe a more regular cellular surveillance system. In either case, it would be gene regulation on the level of epigenesis (Strohman, 1994), which is beyond the level of regulation imposed by a MAR-mediated DNA loop. To verify the apparent dosage compensation, we are investigating the relationship between copy numbers and the resulting expression levels by more classical genetics.

In the LM5(NG) population, higher copy numbers yielded significantly lower GUS activities, despite the large variation in the subpopulations (Figure 6 and Table 2). This inverse correlation in the LM5(NG) population may be due to one or a combination of the various gene-silencing phenomena described for transgenes in plants (Finnegan and McElroy, 1994; Flavell, 1994). If so, the distribution of GUS gene expression in the LM9(ANGA) population indicates that the A element–generated loop may circumvent the occurrence of such silencing phenomena. A similar link between silencing and MAR protection was suggested by Allen et al. (1993), based on GUS expression data in tobacco calli.

Establishing a Genomic Unit of Transcriptional Regulation

The DNA loop hypothesis predicts that MAR elements generate a loop of DNA that establishes a unit of transcriptional regulation (Laemmli et al., 1992). The scattergram of GUS activity against the amount of NPTII in the case of the LM5(NG) population (Figure 5) suggests that the position effects exerted

Table 2. Statistics of Subpopulations of Plants Based on GUS Gene Copy Numbers

	GUS Gene Copy Number[a]														
	1 Copy					2 Copies					> 2 Copies				
Plants	No.[b]	Mean[c]	Var.[d]	F[e]	P[f]	No.	Mean	Var.	F	P	No.	Mean	Var.	F	P
LM5(NG)	27	3.87	0.75	—[g]	—	13	2.73	1.45	—	—	10	2.59	0.80	—	—
LM8(NAGA)	20	4.00	0.66	NS[h]	NA[i]	5	3.96	0.52	NS	NA	2	3.11	1.95	NS	NA
LM9(ANGA)	27	4.23	0.17	4	***[j]	8	4.17	0.04	36	***	12	4.10	0.41	NS	NA

[a] The number of GUS gene copies was taken as the number of left border fragments.
[b] Number of transformants in the subpopulation.
[c] Means are based on the natural logarithms of the GUS activities.
[d] Variance.
[e] Fold reduction of variance with respect to the corresponding LM5(NG) population.
[f] Probability according to the *F* test for homogeneity of variances (Sokal and Rohlf, 1981) compared with the corresponding control population LM5(NG).
[g] (—), control population.
[h] Not significant (at $P = 0.05$).
[i] Not applicable.
[j] Indication of probability value; ***, significant at $P < 0.001$.

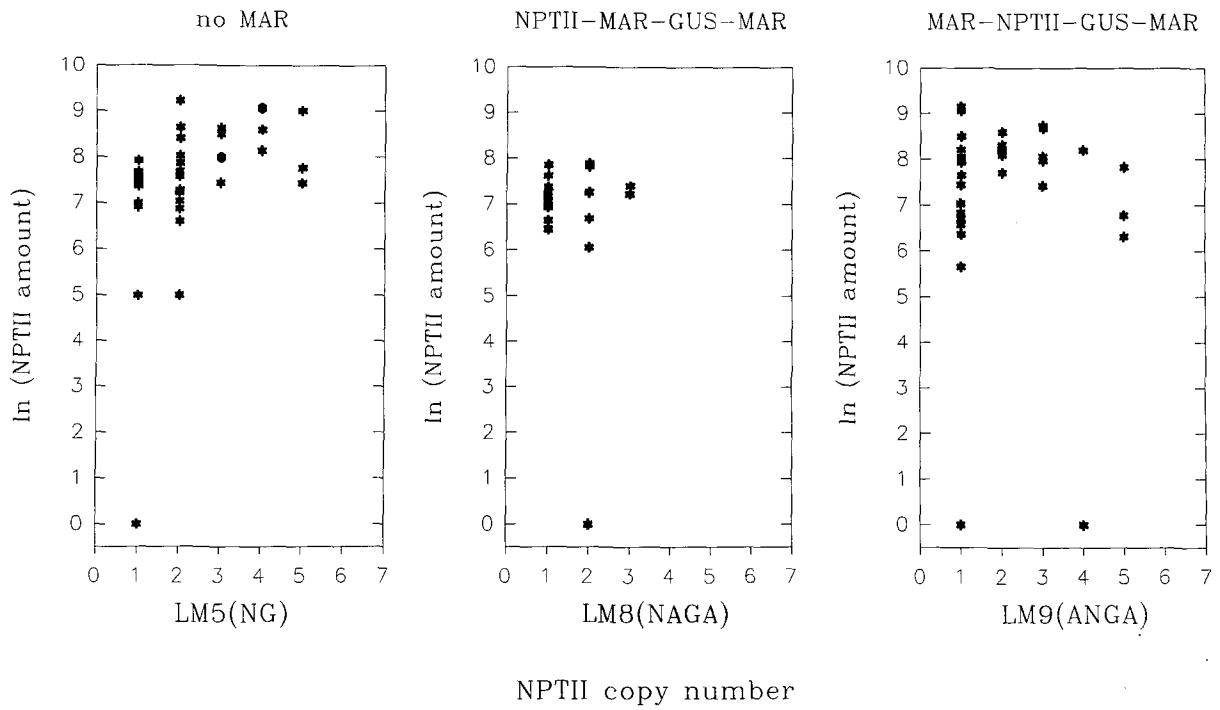

Figure 7. Amount of NPTII as a Function of NPTII Gene Copy Number.

For each of the three populations of plants, the number of NPTII gene copies, as determined by DNA gel blot analysis, is plotted against the natural logarithm (ln) of the amount of NPTII protein. Due to the way these data are plotted, not all points corresponding to individual plants are discernible. For each population, the T-DNA configuration is given above the graphs. The statistics describing the subpopulations of plants are given in Table 3. Linear regression analysis was performed to fit the equation ln(NPTII amount) = $C + r \times$ ln(copy), with regression parameters C (a constant estimating the amount of NPTII) and r (the slope giving the correlation between the amount of NPTII and NPTII gene copy number). This analysis yielded for the LM5(NG) population a C value of 7.3 (±0.3) and an r value of 0.6 (±0.5), which was significant at $P < 0.01$; for the LM8(NAGA) population, this yielded a C value of 7.1 (±0.3) and an r value of 0.3 (±0.6), which was not significant at $P = 0.05$; and for the LM9(ANGA) population, this yielded a C value of 7.7 (±0.4) and an r value of 0.2 (±0.5), which was also not significant at $P = 0.05$.

on each of these genes are independent. Theoretically, the GUS and NPTII activities of all plants with the same copy number should cluster into a single point when position effects are fully eliminated. The results with the LM9(ANGA) plants showed that only the variation in GUS gene activity was reduced. Although it cannot yet be formally excluded that the nopaline synthase (nos) promoter is an exception with respect to the reduction in variability, our results obtained with the unrelated dCaMV 35S and Lhca3.St.1 promoters, as well as results obtained with the nos promoter in callus (Breyne et al., 1992a), suggest that the phenomenon of reduced position effect is independent of the promoter sequence. If so, the lack of reduction in the variability of NPTII gene expression shows that the presence of genes on one and the same loop is no guarantee of MAR efficacy. However, both the presence of a selection bias for NPTII activity and the distance to the A element must be taken into account to explain the lack of a reduction in the variability of the amount of NPTII.

Selection Bias for NPTII Gene Expression

The variation in the amount of NPTII is relatively low in the LM5(NG) population (Table 1), possibly due to a selection bias for NPTII gene expression. Kanamycin selection during transformation requires that shoots express the NPTII gene above a certain threshold level. This creates a selection bias against plants with low NPTII levels. It is therefore conceivable that a putative MAR-mediated reduction of variation in NPTII expression is obscured by selection against low expressers in the LM5(NG) population. Such low expressers would not be present in the LM9(ANGA) transformation due to the supposed MAR influence on the expression of the NPTII gene. As a result, populations would not appear to differ in variation. The higher relative transformation efficiency observed for pLM9(ANGA) (Figure 2) further supports the notion that the A element is influencing NPTII gene expression. Despite the obvious selection bias, the remaining variation above the threshold level

of NPTII gene expression is considerable (Figure 7), indicating that selection bias alone cannot fully account for the lack of a reduction in the variability of NPTII gene expression.

Physical Distance to the MAR Element

Previously, we hypothesized that the size of the MAR-generated loop was important for MAR efficacy. However, when considering all four A element–containing T-DNA vectors evaluated in transgenic plants to date (see also Mlynárová et al., 1994), the physical distance between the promoter and the A element becomes an additional explanatory variable. Whenever the MAR element is physically near a promoter, that is, Lhca3.St.1, dCaMV 35S, or nos, the gene controlled by such a promoter is less subject to a reduction in variability of expression. The affinity of the A element for the nuclear matrix may force a promoter to be too close to the matrix for optimal functioning. Alternatively, the A element sequence may interfere with transcription. The 3′ region of the A element used in our constructs, a 660-bp HaeII-SacI fragment, consists of DNA with a strongly curved structure with almost no affinity for the nuclear matrix (von Kries et al., 1990). It has been hypothesized that this curved DNA guides the DNA away from the nuclear matrix or serves as a binding site for nucleosomes (von Kries et al., 1990). Possibly this curved DNA interacts with the nearby promoter. This interaction may interfere with transcription factor binding and/or formation of an active transcription complex, resulting in unpredictable transcriptional activity. Topological interactions between MARs and promoters have been concluded from animal in vitro studies as well (Schlake et al., 1994).

An empirical indication that the distance to the A element is important comes from the analyses of plant populations in which the A element is immediately adjacent to a promoter. For the 2-kb Lhca3.St.1 promoter (population carrying the pBA T-DNA vector; see Mlynárová et al., 1994), the overall reduction in position effect variability was threefold. For the 0.8-kb dCaMV 35S promoter, this was reduced to a twofold reduction (LM8(NAGA) population; Table 1). No clear reduction was observed for the 0.3-kb nos promoter (LM9(ANGA) population; Table 1), even after taking the selection bias into account (see previous discussion). Currently, we are planning new T-DNA vectors that will allow a more direct evaluation of the role of the curved DNA in transgenic plants.

Boundary Elements, Chromatin Dynamics, and Plant Transgene Expression

In various systems, a great many different boundary elements have been characterized in recent years. All boundary elements influence the expression of genes located nearby but differ in characteristics. Therefore, it is becoming clear that different types of boundaries exist, and these may have different consequences for the regulation of gene expression (Lewin, 1994). The A element as MAR sequence is considered to anchor DNA mechanically to the nuclear matrix; as such it is only one of the various boundary elements now known. Taking the complexity of chromatin dynamics into account (Laemmli et al., 1992), it is surprising that one such element has so significant an influence on the distribution of gene expression in mature transgenic tobacco plants. This is more surprising in view of recent experiments using the full chicken lysozyme locus in transgenic mice; the results of these experiments failed

Table 3. Statistics of Subpopulations of Plants Based on NPTII Gene Copy Numbers

	NPTII Gene Copy Number[a]													
	1 Copy					2 Copies					> 2 Copies			
Plants	No.[b]	Mean[c]	Var.[d]	F[e]	P[f]	No.	Mean	Var.	F	P	No.	Mean	Var.	F	P
LM5(NG)	13	7.22	0.53	—[g]	—	14	7.48	1.03	—	—	10	8.26	0.36	—	—
LM8(NAGA)	11	7.08	0.20	NS[h]	NA[i]	7	7.26	0.47	NS	NA	2	7.32	0.02	NS	NA
LM9(ANGA)	18	7.63	0.99	NS	NA	6	8.20	0.09	11	*[j]	9	7.78	0.67	NS	NA

[a] The number of NPTII gene copies was taken as the number of right border fragments.
[b] Number of transformants in the subpopulation.
[c] Means are based on the natural logarithms of the amounts of NPTII protein.
[d] Variance.
[e] Fold reduction of variance with respect to the corresponding LM5(NG) population.
[f] Probability according to the *F* test for homogeneity of variances (Sokal and Rohlf, 1981) compared with the corresponding control population LM5(NG).
[g] (—), control population.
[h] Not significant (at $P = 0.05$).
[i] Not applicable.
[j] Indication of probability value; *, significant at $P < 0.05$.

to show clear chromatin organizing activity of the very same A element (Huber et al., 1994). A putative role of the A element in suppressing ectopic expression is suggested (Bonifer et al., 1994). Despite all the unknowns with respect to the precise molecular mechanism underlying A element action, the action itself in reducing position effects on transgene expression in transgenic tobacco plants is clear. It, therefore, seems attractive to use the type of T-DNA vector described here for practical applications. Moreover, the relative ease by which transgenic plants can be obtained and analyzed both molecularly and genetically makes plants attractive for the continued study of the action of such boundary elements in a mature organism. As shown in this study, more delicate mechanisms of gene regulation may be hidden underneath the positional variation in transgene expression; these include topological constraints for promoter action or gene dosage compensation.

METHODS

Plant Material and Transformation

Tobacco (*Nicotiana tabacum* cv Petit Havana SR1) plants were grown and transformed with *Agrobacterium tumefaciens* using the leaf disc transformation procedure (Horsch et al., 1985) as described previously (Mlynárová et al., 1994). To ensure independent transformation events, no more than one shoot was harvested from each side of the leaf discs. Selection during transformation was with 50 μg/mL kanamycin, and excised shoots that rooted in the presence of this concentration of kanamycin were considered transgenic.

T-DNA Vector Construction

To obtain pLM5(NG), the cauliflower mosaic virus (CaMV) 35S promoter from pRT103 (Töpfer et al., 1987) was first doubled by cloning the blunt-ended HindIII-EcoRV enhancer fragment in the HincII site. Subsequently, the double CaMV (dCAMV) 35S promoter as a HindIII-NcoI fragment and the β-glucuronidase (GUS) gene–nopaline synthase (nos) terminator from pPATGUS16 (Nap et al., 1992) as an NcoI-EcoRI fragment were ligated into HindIII-, EcoRI-digested pBluescript SK+ (Stratagene). The dCaMV 35S promoter–GUS gene fusion was subsequently cloned as an SalI-BamHI fragment into pBIN19 (Bevan, 1984) to yield pLM5(NG). To obtain pLM8(NAGA), the A element from pUC-B-1-X1 (Phi-Van and Strätling, 1988) was isolated as a BamHI-XbaI fragment, blunt ended, and cloned into the SmaI site of pUC19, and the desired orientation was selected. Subsequently, the SalI-BamHI promoter–GUS gene fragment from pLM5 was cloned upstream of the A element. In this manner, an SalI-Asp718I fragment was generated that carried the dCAMV 35S promoter–GUS–A element combination. This SalI-Asp718I fragment was cloned together with the A element as a BamHI-SalI fragment from pUC-B-1-X1 in BamHI-, Asp718I-digested pBIN19 to generate pLM8(NAGA). The same SalI-Asp718I fragment was cloned in the pBIN19 derivative carrying the A element upstream of the neomycin phosphotransferase (NPTII) gene that was described previously (Mlynárová et al., 1994) to yield pLM9(ANGA). The overall characteristics of the three plant transformation vectors are depicted in Figure 1. Standard cloning techniques (Sambrook et al., 1989) were used in all steps. Binary plasmids were conjugated to *A. tumefaciens* LBA4404 in a biparental mating as described previously (Mlynárová et al., 1994).

Quantitative Gene Assays and Statistics

Fluorometric GUS assays of both in vitro– and greenhouse-grown plants were performed as described previously (Mlynárová et al., 1994). Each plant was assayed four times. GUS activity is expressed as picomoles of methylumbelliferone per minute per microgram of soluble protein. The amount of NPTII in in vitro–grown plant material was determined using an NPTII-ELISA kit (5 Prime → 3 Prime Inc., Boulder, CO). This kit was used essentially according to the manufacturer's recommendations and published modifications (Nagel et al., 1992), employing 50 μg of soluble protein per assay. Each plant was assayed four times, with the amounts of NPTII expressed as femtogram of NPTII protein per microgram of soluble protein. Analyses of the expression data are based on the natural logarithm of the respective activity to give a normal distribution (Nap et al., 1993a). Statistical analyses were performed with the program Genstat 5 (Payne et al., 1987, 1990).

T-DNA Copy Number Determination

All procedures, restriction enzymes, and probes used were as described previously (Mlynárová et al., 1994).

ACKNOWLEDGMENTS

We thank the CPRO-DLO greenhouse team for excellent plant care; Andy Pereira for advice and pointing out the link with gene dosage in maize; and Michiel van Lookeren Campagne, Ruud de Maagd, and Dirk Bosch for useful comments on earlier versions of this manuscript. This research was supported by Grant No. CT910298 from the European Economic Union BRIDGE program. A.J.C. was a collaborator via a fellowship under the Organization for Economic Cooperation and Development project on Biological Resource Management; Ľ.M. was a recipient of grants from the Dutch Ministry of Agriculture and the Dutch Organization for Scientific Research (NWO) while on leave from the Institute of Plant Genetics, Nitra, Slovak Republic.

Received December 27, 1994; accepted March 14, 1995.

REFERENCES

Allen, G.C., Hall, G.E., Jr., Childs, L.C., Weissinger, A.K., Spiker, S., and Thompson, W.F. (1993). Scaffold attachment regions increase reporter gene expression in stably transformed plant cells. Plant Cell **5**, 603–613.

Assaad, F.F., Tucker, K.L., and Signer, E.R. (1993). Epigenetic repeat-induced gene silencing (RIGS) in Arabidopsis. Plant Mol. Biol. **22**, 1067–1085.

Bevan, M. (1984). Binary *Agrobacterium* vectors for plant transformation. Nucleic Acids Res. **12**, 8711–8721.

Birchler, J.A. (1981). The genetic basis of dosage compensation of alcohol dehydrogenase-1 in maize. Genetics **97**, 625–637.

Bonifer, C., Vidal, F., Grosveld, F., and Sippel, A.E. (1990). Tissue specific and position independent expression of the complete gene domain for chicken lysozyme in transgenic mice. EMBO J. **9**, 2843–2848.

Bonifer, C., Yannoutsos, N., Krüger, G., Grosveld, F., and Sippel, A.E. (1994). Dissection of the locus control function located on the chicken lysozyme gene domain in transgenic mice. Nucleic Acids Res. **22**, 4202–4210.

Breyne, P., Van Montagu, M., Depicker, A., and Gheysen, G. (1992a). Characterization of a plant scaffold attachment region in a DNA fragment that normalizes transgene expression in tobacco. Plant Cell **4**, 463–471.

Breyne, P., Gheysen, G., Jacobs, A., Van Montagu, M., and Depicker, A. (1992b). Effect of T-DNA configuration on transgene expression. Mol. Gen. Genet. **235**, 389–396.

Chuang, P.-T., Albertson, D.G., and Meyer, B.J. (1994). DPY-27: A chromosome condensation protein homolog that regulates *C. elegans* dosage compensation through association with the X chromosome. Cell **79**, 459–474.

Dean, C., Jones, J., Favreau, M., Dunsmuir, P., and Bedbrook, J. (1988). Influence of flanking sequences on variability in expression levels of an introduced gene in transgenic tobacco plants. Nucleic Acids Res. **16**, 9267–9283.

Dorer, D.R., and Henikoff, S. (1994). Expansions of transgene repeats cause heterochromatin formation and gene silencing in Drosophila. Cell **77**, 993–1002.

Finnegan, J., and McElroy, D. (1994). Transgene inactivation: Plants fight back! Bio/Technology **12**, 883–888.

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

Hobbs, S.L.A., Warkentin, T.D., and DeLong, C.M.O. (1993). Transgene copy number can be positively or negatively associated with transgene expression. Plant Mol. Biol. **21**, 17–26.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G., and Fraley, R.T. (1985). A simple and general method for transferring genes into plants. Science **227**, 1129–1231.

Huber, M., Bosch, F.X., Sippel, A.E., and Bonifer, C. (1994). Chromosomal position effects in chicken lysozyme gene transgenic mice are correlated with suppression of DNase I hypersensitive site formation. Nucleic Acids Res. **22**, 4195–4201.

Kooter, J.M., and Mol, J.N.M. (1993). *trans*-Inactivation of gene expression in plants. Curr. Opin. Biotechnol. **4**, 166–171.

Laemmli, U.K., Käs, E., Poljak, L., and Adachi, Y. (1992). Scaffold-associated regions: *cis*-Acting determinants of chromatin structural loops and functional domains. Curr. Opin. Genet. Dev. **2**, 275–285.

Lewin, B. (1994). Chromatin and gene expression: Constant questions, but changing answers. Cell **79**, 397–406.

Mlynárová, Ľ., Loonen, A., Heldens, J., Jansen, R.C., Keizer, P., Stiekema, W.J., and Nap, J.P. (1994). Reduced position effect in mature transgenic plants conferred by the chicken lysozyme matrix-associated region. Plant Cell **6**, 417–426.

Nagel, R.J., Manners, J.M., and Birch, R.G. (1992). Evaluation of an ELISA assay for rapid detection and quantification of neomycin phosphotransferase II in transgenic plants. Plant Mol. Biol. Rep. **10**, 263–272.

Nap, J.-P., Dirkse, W.G., Louwerse, J., Onstenk, J., Visser, R., Loonen, A., Heidekamp, F., and Stiekema, W.J. (1992). Analysis of the region in between two closely linked patatin genes: Class II promoter activity in tuber, root and leaf. Plant Mol. Biol. **20**, 683–694.

Nap, J.-P., Keizer, P., and Jansen, R. (1993a). First-generation transgenic plants and statistics. Plant Mol. Biol. Rep. **11**, 156–164.

Nap, J.-P., van Spanje, M., Dirkse, W.J., Baarda, G., Mlynárová, Ľ., Loonen, A., Grondhuis, P., and Stiekema, W.J. (1993b). Activity of the promoter of the *Lhca3.St.1* gene, encoding the potato apoprotein 2 of the light-harvesting complex of photosystem I, in transgenic potato and tobacco plants. Plant Mol. Biol. **23**, 605–612.

Payne, R.W., Lane, P.W., Ainsley, A.E., Bicknell, K.E., Digby, P.G.N., Harding, S.A., Leech, P.K., Simpson, H.R., Todd, A.D., Verrier, P.J., White, R.P., Gower, J.C., Tunnicliffe Wilson, G., and Paterson, L.J. (1987). Genstat 5 Reference Manual. (Oxford, UK: Clarendon Press).

Payne, R.W., Lane, P.W., Ainsley, A.E., Bicknell, K.E., Digby, P.G.N., Harding, S.A., Leech, P.K., Simpson, H.R., Todd, A.D., Welham, S.J., White, R.P., Thompson, R., Tunnicliffe Wilson, G., Underhill, L.G., and Verrier, P.J. (1990). Genstat 5 (Release 2) Reference Manual Supplement. (Oxford, UK: Numerical Algorithms Group Ltd.).

Peach, C., and Velten, J. (1991). Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. Plant Mol. Biol. **17**, 49–60.

Phi-Van, L., and Strätling, W.H. (1988). The matrix attachment regions of the chicken lysozyme gene co-map with the boundaries of the chromatin domain. EMBO J. **7**, 655–664.

Reitman, M., Lee, E., Westphal, H., and Felsenfeld, G. (1990). Site-independent expression of the chicken β^A-globin gene in transgenic mice. Nature **348**, 749–752.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, 2nd ed. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Schedl, A., Montoliu, L., Kelsey, G., and Schütz, G. (1993). A yeast artificial chromosome covering the tyrosinase gene confers copy number–dependent expression in transgenic mice. Nature **362**, 258–261.

Schlake, T., Klehr-Wirth, D., Yoshida, M., Beppu, T., and Bode, J. (1994). Gene expression within a chromatin domain: The role of core histone hyperacetylation. Biochemistry **33**, 4197–4206.

Smith, H.A., Swaney, S.L., Parks, T.D., Wernsman, E.A., and Dougherty, W.G. (1994). Transgenic plant virus resistance mediated by untranslatable sense RNAs: Expression, regulation, and fate of nonessential RNAs. Plant Cell **6**, 1441–1453.

Sokal, R.R., and Rohlf, F.J. (1981). Biometry, 2nd ed. (New York: Freeman and Co.).

Stief, A., Winter, D.M., Strätling, W.H., and Sippel, A.E. (1989). A nuclear DNA attachment element mediated elevated and position-independent gene activity. Nature **341**, 343–345.

Strohman, R. (1994). Epigenesis: The missing beat in biotechnology? Bio/Technology **12**, 156–164.

Töpfer, R., Matzeit, V., Gronenborn, B., Schell, J., and Steinbiss, H.-H. (1987). A set of plant expression vectors for transcriptional and translational fusions. Nucleic Acids Res. **15**, 5890.

van der Geest, A.H.M., Hall, G.E., Jr., Spiker, S., and Hall, T.C. (1994). The β-phaseolin gene is flanked by matrix attachment regions. Plant J. **6**, 413–423.

von Kries, J.P., Phi-Van, L., Diekmann, S., and Strätling, W.H. (1990). A non-curved chicken lysozyme 5′ matrix attachment site is 3′ followed by a strongly curved DNA sequence. Nucleic Acids Res. **18**, 3881–3885.