The Plant Cell, Vol. 8, 1589–1599, September 1996

Approaching the Lower Limits of Transgene Variability

Ľudmila Mlynárová,[a] L. C. Paul Keizer,[b] Willem J. Stiekema,[a] and Jan-Peter Nap[a,1]

[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

The inclusion of chicken lysozyme matrix-associated regions (MARs) in T-DNA has been demonstrated to reduce the variation in β-glucuronidase (GUS) gene expression among first-generation transformed plants. The residual variation observed between transgenic plant lines with MARs at the T-DNA borders was investigated. By definition, any phenotypic variance between or within genetically identical plants is caused by random or environmental variation. This variation therefore sets a lower limit to the variation in GUS activities. The variance of GUS activity in offspring plant populations of genetically identical individuals was used as an estimate of environmental variation. For transgenic plants with MARs at the T-DNA borders, the variation between independent transformants could not be distinguished from the environmental variation. The variation could be attributed mainly to the variation in the GUS activity measurement. Therefore, the MAR element approached the maximal possible reduction of transgene variability given current technology and sample sizes. The role of MARs in offspring plants was evaluated by comparing such populations of transgenic plants for the magnitude of and variation in GUS activity. Pairwise comparisons showed that the presence of MARs reduced variation in offspring generations in the same manner as demonstrated for primary transformants. The populations carrying a doubled cauliflower mosaic virus 35S promoter-GUS gene tended to be more variable than the *Lhca3.St.1* promoter-GUS gene–carrying populations. This tendency indicated an intrinsic susceptibility of the doubled cauliflower mosaic virus 35S promoter to variation. Homozygous plants were approximately twice as active as the corresponding hemizygous plants and tended to be more variable than the hemizygous plants. We hypothesized that the magnitude of environmental variations is related to a higher susceptibility to transgene silencing.

INTRODUCTION

Genetic transformation of plants generally results in a large and seemingly random variation in the expression of the newly introduced transgene between individual transformants (Peach and Velten, 1991; Nap et al., 1993a). This variability is attributed to different integration sites of the transgene, reflecting the influences of the surrounding chromatin known as position effects. However, it is becoming clear that differences in transgene copy number and transgene configuration, notably repeat configurations (Hobbs et al., 1993) and various (epigenetic) silencing phenomena (Matzke and Matzke, 1993; Finnegan and McElroy, 1994; Flavell, 1994), also contribute to the observed variability in transgene expression. Variation in transgene expression levels is undesirable for applications of transgenic plants in science and product development as well as from regulatory perspectives. It reduces the predictability and efficiency of genetic transformation and necessitates the establishment and analysis of many independently selected transformants to obtain the desired phenotype with the appropriate stability of transgene expression (Conner and Christey, 1994). In addition, further insight into the causes of the variability of transgene expression is likely to contribute to our understanding of the regulation of plant gene expression.

Our approach to reduce the variability of transgene expression is the inclusion of chromatin boundary elements at the borders of the Agrobacterium T-DNA (Mlynárová et al., 1994, 1995). One class of boundary elements, the matrix-associated regions (MARs), is thought to insulate genes from the influences of the surrounding chromatin (Laemmli et al., 1992) or to prevent mislocalization of genes in the heterochromatin (Dorer and Henikoff, 1994) by virtue of their affinity for the nuclear matrix. The boundary element used in our studies is the chicken lysozyme MAR known as the A element (Stief et al., 1989). Flanking transgenes with A elements at the borders of the Agrobacterium T-DNA significantly reduced variability in β-glucuronidase (GUS) gene expression in mature transgenic tobacco plants, irrespective of the promoter sequence used (Mlynárová et al., 1994, 1995). Important questions to be addressed are as follows: (1) What is the lower limit to which the variability in transgene expression could be reduced? (2) Do

[1] To whom correspondence should be addressed.

MAR sequences reduce variability in transgene expression in subsequent generations?

By definition, any difference in phenotypic variance between or within genetically identical plants is due to differences in the environmental component of the variation (Falconer, 1981). This so-called environmental variation may be composed of several factors, such as experimental errors, macroenvironmental deviations, microenvironmental deviations, and/or developmental differences. Such environmental variation is essentially random in nature and cannot be completely controlled. Therefore, the environmental variation may provide an estimate for the lowest variation attainable for a particular transgene. It establishes the lower limit that can be attained for the reduction in variability of transgene expression. Comparing the environmental variation with the variation observed in the populations of primary transformants will indicate how much further transgene variability can be reduced theoretically. In this report, we present an analysis of GUS transgene expression in large populations of genetically identical plants. The results establish that the variability between individual, independently selected MAR-carrying transformants could not be distinguished from the variability of genetically homogeneous individuals carrying the MAR-contained GUS gene. The A element therefore accomplished a reduction of transgene variability that approaches the limits set by current technology and sample sizes.

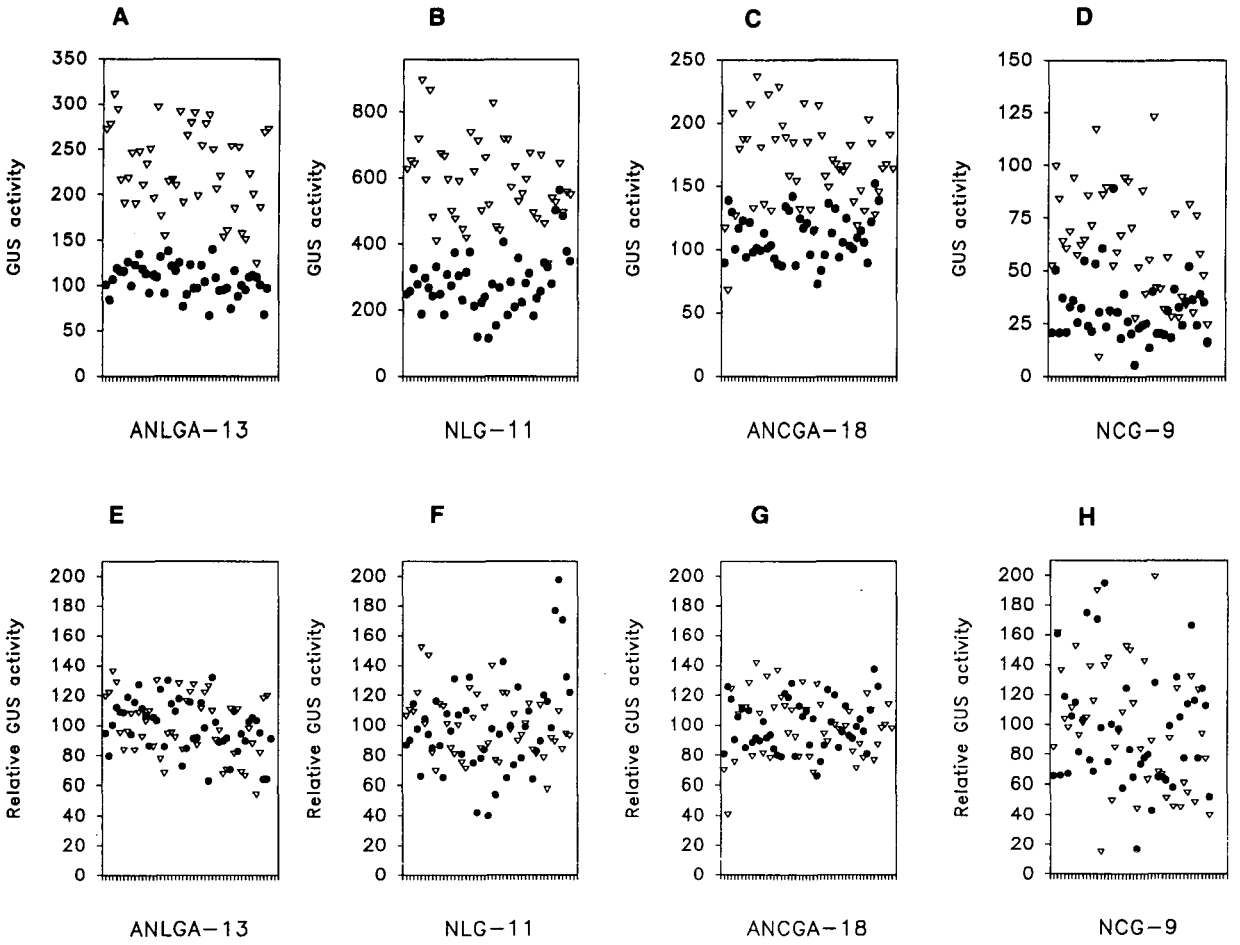

Figure 1. Distribution of Absolute and Relative GUS Activities in Eight Offspring Populations of Genetically Identical Plants.

(A) to **(D)** Absolute GUS activity is plotted in picomoles of 4-methylumbelliferone per minute per microgram of soluble protein.

(E) to **(H)** Relative GUS activity is plotted as a percentage of the mean of that population, which is set at 100.

In **(A)** to **(H)**, each graph shows two populations that have identical T-DNA but differ in the zygosity of the T-DNA. In all of the graphs, plants carrying the T-DNA in the homozygous state are represented by open triangles. The corresponding hemizygous populations are represented by closed circles. Each point on the x-axis represents an individual plant. The names of the populations are given below each graph. Details of these populations are given in Table 1. The statistics accompanying these populations are given in Table 2.

RESULTS

Statistics of the Offspring Populations

GUS gene activity in all individuals of the offspring populations from eight independently derived transgenic plants is shown in Figures 1A to 1D. The characteristics of these eight populations are given in Table 1, and the overall statistics describing them are presented in Table 2. GUS activities were evaluated with the Lillifors normality test (Conover, 1980; Nap et al., 1993a). GUS activity in the homozygous (Ho) and hemizygous (He) offspring plants was normally distributed, with the slight exception of the NCG-9-He population (Table 2). The variation in GUS activity strongly depended on the mean (or median), which is illustrated in Figures 1A to 1D and in the distribution of variances given in Table 2. As a result of such a relationship between mean and variance, the variability of the GUS measurement depended on the absolute activity measured.

The dependence of variance on the mean could indicate a multiplicative effect either in GUS gene expression or in the fluorometric determination of GUS activity. To enable us to compare such measurements, a measure of relative variability was required. The distribution of GUS activity in all eight populations as a percentage of the mean for each population is illustrated in Figures 1E to 1H, with the accompanying coefficients of variation (CV) given in Table 2. In six of the eight populations (ANLGA-13-Ho and ANLGA-13-He, NLG-11-Ho and NLG-11-He, and ANCGA-18-Ho and ANCGA-18-He), the CV values are of the same order of magnitude. However, the CV values of the NCG-9-Ho and NCG-9-He populations were higher (Table 2). Previously, we have shown that a logarithmic transformation is required to yield an approximately normal distribution of GUS gene activity in a population of primary transformants (Nap et al., 1993a). The variance and coefficient of variation on the logarithmic scale are given in Table 2. Upon this logarithmic transformation, the Lillifors test indicates that all eight populations can reasonably be assumed to follow a normal distribution (Table 2), allowing all statistical analyses that require such a normal distribution.

Table 1. Description of the Populations of Plants Used in This Study

Population Name[a]	T-DNA[b]	MAR[c]	Promoter	Zygosity[d]
Offspring plants, large populations				
ANLGA-13-Ho	pLM	+	*Lhca3.St.1*	Homo
ANLGA-13-He	pLM	+	*Lhca3.St.1*	Hemi
NLG-11-Ho	pPPG	–	*Lhca3.St.1*	Homo
NLG-11-He	pPPG	–	*Lhca3.St.1*	Hemi
ANCGA-18-Ho	pLM9(ANGA)	+	dCaMV	Homo
ANCGA-18-He	pLM9(ANGA)	+	dCaMV	Hemi
NCG-9-Ho	pLM5(NG)	–	dCaMV	Homo
NCG-9-He	pLM5(NG)	–	dCaMV	Hemi
Offspring plants, small populations				
NLG-1-He	pPPG	–	*Lhca3.St.1*	Hemi
NLG-10-He	pPPG	–	*Lhca3.St.1*	Hemi
NLG-15-He	pPPG	–	*Lhca3.St.1*	Hemi
NLG-41-He	pPPG	–	*Lhca3.St.1*	Hemi
NCG-4-He	pLM5(NG)	–	dCaMV	Hemi
NCG-10-He	pLM5(NG)	–	dCaMV	Hemi
NCG-47-He	pLM5(NG)	–	dCaMV	Hemi
NCG-52-He	pLM5(NG)	–	dCaMV	Hemi
First-generation transformants				
LM(ANLGA)-all	pLM	+	*Lhca3.St.1*	Hemi
LM9(ANCGA)-all	pLM9(ANGA)	+	dCaMV	Hemi
LM(ANLGA)-1copy	pLM	+	*Lhca3.St.1*	Hemi
LM9(ANCGA)-1copy	pLM9(ANGA)	+	dCaMV	Hemi

[a] The name of the population of offspring plants is an acronym of the T-DNA configuration carried by the plants in the population. A represents the chicken lysozyme MAR element; N, the NPTH gene; L, the *Lhca3.St.1* promoter; C, the dCaMV 35S promoter; G, the GUS gene; Ho, homozygous; and He, hemizygous. In addition, the number identifies the original primary transformant. The name of the population of first-generation transformants refers to the T-DNA vector that the plants carry. The suffix "-all" indicates that it concerns all transformants; the suffix "-1copy" indicates that it concerns the subset of plants carrying one intact copy of the T-DNA.

[b] Transformation vector as described in the text and in Mlynárová et al. (1994, 1995).

[c] Presence (+) or absence (–) of the chicken lysozyme MAR element at the T-DNA borders.

[d] Homo, homozygous individuals; hemi, hemizygous individuals.

Putative Outliers

A method to evaluate the individual observations in populations is half-normal plot analysis, in which standardized residuals are plotted against normal distribution statistics (Lane and Payne, 1994; Sokal and Rohlf, 1995). Such a plot should yield a straight line in the case of a normal distribution. Inspection of the half-normal plots belonging to these eight populations indicated that one plant in the NCG-9-Ho population and two plants in the NCG-9-He population should be considered outliers (plots not shown). We had no biological reasons to discard these plants from the analyses. To assess the relative importance of these plants, we removed their data points from the populations, yielding the populations NCG-9-Ho-t1 and NCG-9-He-t2. The descriptive statistics for these two populations are also given in Table 2. This procedure established that the variance and coefficient of variation of these NCG populations remained relatively high, even when the outliers were omitted.

Table 2. Overall Statistics of the Large Offspring Populations

		Scale of Measurement						Natural Logarithmic Scale					
Population	No.[a]	Mean[b]	SEM	Var[c]	CV[d]	Med[e]	P[f]	Mean	SEM	Var	CV	Med	P
ANLGA-13-Ho	45	227.56	7.02	2219.49	20.7	220.72	+	5.41	0.033	0.048	4.1	5.40	+
ANLGA-13-He	45	105.82	2.61	305.95	16.5	108.04	+	4.65	0.026	0.030	3.8	4.68	+
NLG-11-Ho	45	589.38	18.13	14785.1	20.6	591.81	+	6.36	0.031	0.043	3.2	6.38	+
NLG-11-He	45	284.62	13.72	8469.74	32.3	276.87	+	5.60	0.049	0.108	5.9	5.62	+
ANCGA-18-Ho	48	161.98	5.02	1210.59	20.8	166.86	+	5.09	0.033	0.053	3.6	5.12	+
ANCGA-18-He	44	110.50	2.78	339.59	16.7	107.68	+	4.69	0.025	0.028	3.6	4.68	+
NCG-9-Ho	44	61.70	3.94	681.32	42.3	59.70	+	4.02	0.076	0.255	12.6	4.09	+
NCG-9-Ho-t1[g]	43	62.92	3.83	630.80	39.9	60.59	+	4.06	0.065	0.184	10.6	4.10	+
NCG-9-He	44	31.24	2.23	218.35	47.3	28.15	(+)	3.34	0.070	0.218	14.0	3.34	+
NCG-9-He-t2[h]	42	30.47	1.76	130.32	37.5	28.15	(+)	3.35	0.055	0.129	10.7	3.34	+

[a] No., number of plants in the population.
[b] Mean, GUS activity of the population in picomoles of 4-methylumbelliferone per minute per microgram of soluble protein.
[c] Var, variance.
[d] CV, coefficient of variation in percentage.
[e] Med, median.
[f] P, probability that population can be considered to be normally distributed according to the Lillifors normality test (Conover, 1980). (+), $0.01 < P < 0.05$; +, $P > 0.10$.
[g] t1, population from which one plant was omitted on the basis of half-normal plot analysis.
[h] t2, population from which two plants were omitted on the basis of half-normal plot analysis.

Variation in First-Generation Transformants Compared with Variation in Offspring Plants

In terms of quantitative genetics, GUS activity in a large population of genetically homogeneous plants can be considered to represent the genotypic value of GUS activity for that particular plant line and, as a consequence, for that particular integration locus. We have compared the variability among the first-generation transformants analyzed previously (Mlynárová et al., 1994, 1995) with the variability in the populations of genetically identical plants described here. The characteristics of the full populations of the first-generation transformants are given in Table 1 with the suffix "-all." On the basis of DNA gel blot analysis, subpopulations of plants containing one copy of the T-DNA were defined (Mlynárová et al., 1994, 1995). These are indicated in Table 1 with the suffix "-1copy."

Half-normal plot analysis of these four populations indicated that a few plants in these populations also did not appear to belong to a normally distributed population. This applied to one plant of the LM(ANLGA)-all population, two plants of the LM(ANLGA)-1copy population, and two plants of the LM9(ANCGA)-all population, one of which was a one-copy plant that was also part of the LM9(ANCGA)-1copy population. Detailed DNA gel blot analysis indicated that this aberrant LM9(ANCGA)-1copy plant had a more complex T-DNA integration pattern (data not shown).

To assess the relative importance of these few plants to the variance of the populations, we made comparisons with and without their inclusion. Comparisons between the populations of genetically homogeneous offspring plants and the populations of first-generation transformants are given in Table 3. It can be concluded that the putative outliers exerted a strong influence on the variance of the populations of first-generation transformants. When the outliers were omitted, the ANCGA-18-He population had approximately the same variance as the LM9(ANCGA)-1copy-t1 population. Therefore, the variation in the individual transgenic plants of the LM9(ANCGA)-1copy-t1 population could not be distinguished from the variation among the genetically identical individuals of the ANCGA-18-He population. Likewise, the variation in the LM(ANLGA)-1copy-t2 population approached the variation in the corresponding ANLGA-13-He population.

Variation in the Measurement of GUS Activity

To estimate the variability due to the GUS measurement, one individual plant from each of the ANLGA-13-He and ANCGA-18-He populations was randomly selected. From these two individual plants, one leaf disc was sampled and assayed 46 times. Additional assays using a number of different samples established that the variances obtained were sufficiently reliable estimates for the error in the GUS measurement (data not shown). The variances obtained from the one-sample repetitions were compared with the variances observed in the offspring populations (Table 4). The variances in these paired comparisons could not be distinguished from one another, showing that the variance observed in the ANLGA-13-He and

Table 3. Comparisons of the Variability in Populations of Offspring Plants and Populations of First-Generation Transformed Plants

Offspring Plants			First-Generation Plants					
Population	No.[a]	Var[b]	Population	No.	Var	P[c]	Fold[d]	UCL[e]
ANLGA-13-He	45	0.030	LM(ANLGA)-all	53	0.399	***	13.2	
			LM(ANLGA)-all-t1[f]	52	0.335	***	11.1	
			LM(ANLGA)-1copy	22	0.201	***	6.70	
			LM(ANLGA)-1copy-t2[g]	20	0.049	NS∼	1.62	2.78
ANCGA-18-He	44	0.028	LM9(ANCGA)-all	55	0.185	***	6.59	
			LM9(ANCGA)-all-t2	53	0.068	**	2.42	
			LM9(ANCGA)-1copy	27	0.173	***	6.15	
			LM9(ANCGA)-1copy-t1	26	0.037	NS	1.31	2.13

[a] No., number of plants in the population.
[b] Var, variance of the natural logarithm–transformed GUS activity.
[c] P, probability that the variance in the population of the first-generation transformants is significantly different from the variance in the corresponding offspring populations, according to the *F* test for homogeneity of variances (Sokal and Rohlf, 1995); NS, not significant ($P > 0.10$); NS∼, $0.05 < P < 0.10$; **, $0.05 < P < 0.001$; ***, $P < 0.001$.
[d] Fold, fold difference in variability, expressed as variance ratio, of the population of first-generation plants relative to the offspring plants.
[e] UCL, one-sided 90% upper confidence limit of the variance ratio.
[f] t1, population from which one plant was omitted on the basis of half-normal plot analysis.
[g] t2, population from which two plants were omitted on the basis of half-normal plot analysis.

ANCGA-18-He offspring populations was due mainly to the experimental variability in the GUS measurements.

Comparisons that are not significantly different at a given probability do not necessarily imply that the populations from which the individuals were drawn are identical. The so-called type II or β error (Sokal and Rohlf, 1995) is in this case the probability that the variances are considered identical when in reality they are not. According to the power function of the *F* test (Lehmann, 1986), the β error for the variance ratios of the comparisons in Tables 3 and 4 that are not significant (at $P = 0.05$) is ∼0.7.

Our interest, however, has been to establish how close the variances of the populations are to one another. Therefore, we determined the confidence bounds of the variance ratios according to Neter et al. (1988). In Tables 3 and 4, the 90% upper confidence limits are given for all comparisons that are not significant (at $P = 0.05$). A 90% upper confidence limit indicates that with 90% probability, the true variance ratio is between the calculated upper bound and unity. In our comparisons, the 90% upper confidence limits also ensured reasonable β errors of ∼0.1 to 0.2 (data not shown). When 95% probability is desired, the confidence limits become multiplied by ∼1.15. The 90% upper confidence limits of the variance ratio of the first-generation transformants and the repeated measurements are 2.8 for the LM(ANLGA)-1copy-t2 population and 2.1 for the LM9(ANCGA)-1copy-t1 population.

Table 4. Comparisons of the Variability in Populations of Offspring Plants and Repeated Measurements of a Single Sample

Offspring Plants			Repeated Measurements				
Population	No.[a]	Var[b]	Plant[c]	Rep[d]	Var	P[e]	UCL[f]
ANLGA-13-He	45	0.030	ANLGA-13-He-19	46	0.022	NS	2.05
ANCGA-18-He	44	0.028	ANCGA-18-He-3	46	0.020	NS	2.04

[a] No., number of plants in the population.
[b] Var, variance of the natural logarithm–transformed GUS activity.
[c] Plant, randomly chosen individual from the corresponding offspring population.
[d] Rep, number of repetitions of the measurement.
[e] P, probability that the variance in the offspring population is significantly different from the variance of the repeated measurements, according to the *F* test for homogeneity of variances (Sokal and Rohlf, 1995); NS, not significant ($P > 0.10$).
[f] UCL, one-sided 90% upper confidence limit of the variance ratio.

Table 5. Pairwise Comparisons of GUS Activity Data

MAR Presence					Zygosity					Promoter Type				
With MAR Population 1	No MAR Population 2	Sed[a]	Fold[b]	P[c]	Homozygous Population 1	Hemizygous Population 2	Sed	Fold	P	*Lhca3.St.1* Population 1	dCaMV Population 2	Sed	Fold	P
ANLGA-13-Ho	NLG-11-Ho	0.045	0.39	***	ANLGA-13-Ho	ANLGA-13-He	0.042	2.13	***	ANLGA-13-Ho	ANCGA-18-Ho	0.046	1.38	***
ANLGA-13-He	NLG-11-He	0.055	0.39	***	NLG-11-Ho	NLG-11-He	0.058	2.13	***	ANLGA-13-He	ANCGA-18-He	0.036	1.00	NS
ANCGA-18-Ho	NCG-9-Ho	0.083	2.94	***	ANCGA-18-Ho	ANCGA-18-He	0.042	1.50	***	NLG-11-Ho	NCG-9-Ho	0.082	10.4	***
ANCGA-18-He	NCG-9-He	0.075	3.86	***	NCG-9-Ho	NCG-9-He	0.104	1.96	***	NLG-11-He	NCG-9-He	0.086	9.58	***

[a] Sed, standard error of the difference between the means of the two populations on the natural logarithmic scale, estimated by restricted maximum likelihood analysis (Payne et al., 1993).

[b] Fold, fold difference in activity of the first population relative to the second population on the scale of measurement, calculated by retransforming the difference between the two means determined on the natural logarithmic scale.

[c] P, probability of the difference between the two means being significant according to the Sed values; NS, not significant ($P > 0.10$); ***, $P < 0.001$.

Comparisons of the Magnitude of and Variation in GUS Gene Expression Levels

The eight offspring populations examined differ in three characteristics: (1) the presence of the MAR sequence; (2) zygosity; and (3) the type of promoter. The standard errors of the means of the natural logarithm–transformed GUS activities were estimated by restricted maximum likelihood analysis (Payne et al., 1993) and used for the pairwise comparisons presented in Table 5. Pairwise comparisons using two-sample *t* tests or the nonparametric Wilcoxon-Mann-Whitney test yielded identical results (data not shown). The homozygous populations were approximately twice as active as the corresponding hemizygous populations. In the case of the *Lhca3.St.1* promoter, the non-MAR populations (NLG-11-Ho and NLG-11-He) had a significantly higher activity than the MAR-containing populations (ANLGA-13-Ho and ANLGA-13-He), whereas for the doubled cauliflower mosaic virus 35S (dCaMV) promoter, the reverse was true.

The relative variabilities of the populations analyzed were compared pairwise using the parametric *F* test (Sokal and Rohlf, 1995). These comparisons are presented in Table 6. The non-MAR-containing populations (NLG-11-Ho and NLG-11-He, and NCG-9-Ho and NCG-9-He) exhibited a higher relative variability than did the MAR-containing populations (ANLGA-13-Ho and ANLGA-13-He, and ANCGA-18-Ho and ANCGA-18-He). The populations of MAR-containing hemizygous plants were approximately half as variable as were the corresponding homozygous populations, with the opposite being observed for the non-MAR *Lhca3.St.1* promoter–carrying plants. The non-MAR dCaMV-carrying populations had a higher relative variability than did the corresponding non-MAR *Lhca3.St.1*-carrying

Table 6. Pairwise Comparisons of GUS Variability Data

MAR Presence				Zygosity				Promoter Type			
With MAR Population 1	No MAR Population 2	P[a]	Fold[b]	Homozygous Population 1	Hemizygous Population 2	P	Fold	*Lhca3.St.1* Population 1	dCaMV Population 2	P	Fold
ANLGA-13-Ho	NLG-11-Ho	NS	1.0	ANLGA-13-Ho	ANLGA-13-He	NS~	0.63	ANLGA-13-Ho	ANCGA-18-Ho	NS	1.0
ANLGA-13-He	NLG-11-He	***	3.57	NLG-11-Ho	NLG-11-He	**	2.53	ANLGA-13-He	ANCGA-18-He	NS	1.0
ANCGA-18-Ho	NCG-9-Ho	***	4.86	ANCGA-18-Ho	ANCGA-18-He	*	0.53	NLG-11-Ho	NCG-9-Ho	***	5.99
ANCGA-18-Ho	NCG-9-Ho-t1[c]	***	3.51	NCG-9-Ho	NCG-9-He	NS	1.0	NLG-11-Ho	NCG-9-Ho-t1	***	4.32
ANCGA-18-He	NCG-9-He	***	7.76	NCG-9-Ho-t1	NCG-9-He-t2	NS	1.0	NLG-11-He	NCG-9-He	*	2.02
ANCGA-18-He	NCG-9-He-t2[d]	***	4.59					NLG-11-He	NCG-9-He-t2	NS	1.0

[a] P, probability of the differences between variances of the natural logarithm–transformed GUS activites exhibited by the two populations being significant according to the *F* test for homogeneity of variances (Sokal and Rohlf, 1995); NS, not significant ($P > 0.10$); NS~, $0.05 < P < 0.10$; *, $0.01 < P < 0.05$; **, $0.001 < P < 0.01$; ***, $P < 0.001$.

[b] Fold, fold difference in variability of the second population relative to the first population.

[c] t1, NCG-9-Ho population from which one plant was omitted on the basis of half-normal plot analysis.

[d] t2, NCG-9-He population from which two plants were omitted on the basis of half-normal plot analysis.

populations. This difference was not apparent in the presence of the MAR elements. In the comparisons that were not significant (at P = 0.05), the 90% upper confidence limit of the variance ratio ranged between 1.6 and 2.3.

The NCG-9-Ho and NCG-9-He populations in particular exhibited high relative variability, irrespective of the omission of outlier values (Table 2). This finding may reflect position effects due to the specific locus of integration. To assess the importance of the specific integration locus, small non-MAR-containing populations of six to 12 hemizygous plants comprising four additional one-copy insertion loci were analyzed for the variability of GUS gene expression. The variances are presented in Table 7. A box plot representation of the distribution of these variances is given in Figure 2. Inspection of the variances obtained showed that the NLG-11-He population had the highest variability of the five NLG populations analyzed. In contrast, the NCG-9-He-t2 population is not an extreme among the five NCG populations analyzed. Although based on relatively small numbers of plants and insertion loci, the results suggested that overall, the non-MAR dCaMV-carrying populations exhibited a higher variance and a higher spread in that variance than did the non-MAR *Lhca3.St.1*-carrying populations (Figure 2).

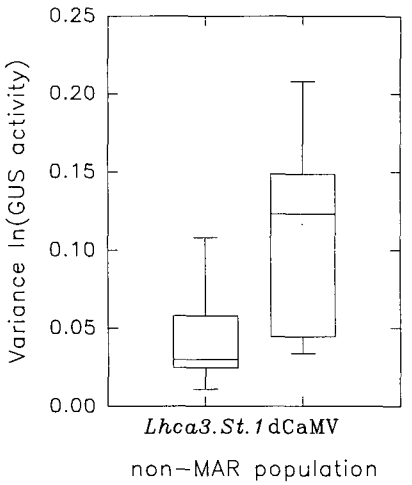

Figure 2. Box Plots of the Distribution of Variances in Populations of Non-MAR-Containing Transformants.

The variances of the natural logarithm–transformed GUS activities of the constituent populations are plotted as box plot representations (Sokal and Rohlf, 1995) in which each horizontal line represents the 10th, 25th, 50th (median), 75th, and 90th percentiles, respectively. The promoter that distinguishes the two non-MAR-containing groups of populations is indicated on the x-axis. The variances of the constituent plant populations are given in Table 7.

DISCUSSION

Genotypic Value and Environmental Variation

Theoretically, quantitative genetics in its most simple form assumes a phenotypic value, for example, GUS activity, to be the sum of a genotypic value G and the environmental deviation E, in which G and E are independent. The genotypic value, by definition, is a theoretical constant to be obtained by evaluating many identical plants, preferably in many locations and under many conditions. The fact that G is a constant implies that it has zero variance. Hence, all variation observed among genetically identical plants is supposed to be of environmental, or "random," origin. Any difference in phenotypic variance between or within inbred plants therefore must be attributed to differences in the environmental component of the variation (Falconer, 1981). This environmental variation is composed of several possible factors, which include experimental error, developmental variation, microenvironmental variation, and others. Such a composed environmental variation will be random, always present, and impossible to eliminate totally. As a consequence, the environmental variance will set the lowest attainable limit of transgene variability. The observation that in most genetically identical offspring populations, GUS activity itself was reasonably normally distributed before a logarithmic transformation (Table 2), already indicated that the nature of the variation was primarily random.

Previously, we demonstrated that the addition of MAR sequences at the borders of the T-DNA reduces variation in transgene expression in populations of primary transformants (Mlynárová et al., 1994, 1995). The experiments described in this study were designed to investigate the proportion of variation in populations of first-generation transformants that could be attributed to the random variation observed in genetically identical offspring populations. Comparisons between such populations indicate how close the observed variation is to the theoretical lower limit of transgene variability. The GUS activities and variances in relatively large populations of ~45 genetically identical individuals of different transgenic plants allowed estimation of the genotypic values of the plant lines

Table 7. Variance in Offspring Populations of Different One-Copy, Hemizygous, Non-MAR-Containing Transformants

Lhca3.St.1 Promoter			dCaMV Promoter		
Population	No.[a]	Var[b]	Population	No.	Var
NLG-1-He	6	0.011	NCG-4-He	9	0.034
NLG-10-He	10	0.041	NCG-10-He	8	0.123
NLG-15-He	7	0.030	NCG-47-He	12	0.208
NLG-41-He	8	0.029	NCG-52-He	8	0.048
NLG-11-He	45	0.108	NCG-9-He-t2[c]	42	0.129

[a] No., number of plants in the population.
[b] Var, variance of the natural logarithm–transformed GUS activities.
[c] NLG-9-He population from which two plants were omitted on the basis of half-normal plot analysis.

chosen as well as the environmental deviations in the eight populations (Table 2). We compared these values with the corresponding values of the populations of first-generation transformants previously obtained. Comparisons among the offspring populations indicated the extent of MAR efficacy.

At the Limits of the Reduction of Transgene Variability

Comparison of the variance of the ANCGA-18-He population with the variance of the LM9(ANCGA)-1copy-t1 population (Table 3) showed that after removal of one putative outlier, the variation in the LM9(ANCGA)-1copy-t1 population was not significantly different (at $P = 0.05$) from the variation in the ANCGA-18-He offspring population. Similarly, the variation in the ANLGA-13-He offspring population was not significantly different (at $P = 0.05$) from the variation in the LM(ANLGA)-1copy-t2 population. Moreover, the variation among the genetically identical offspring plants could not be distinguished from the random error in the GUS measurement (Table 4).

These comparisons were used to establish how close variances come to each other. For this purpose, confidence limits of variance ratios are more appropriate than lack of significant differences. Upper confidence limits indicated with 90% probability that the variance ratio of GUS activity in genetically identical plants and GUS activity in repeated measurements of a single sample was two or less (Table 4). Apparently, we accomplished very homogeneous sampling. Undoubtely there was variation within plants, for example, between leaves of different age or position. However, the results show that such a variation was virtually eliminated in the standardized experimental setup used.

The upper confidence limits indicated that with 90% probability, the variance ratio of GUS activity in genetically identical plants and in first-generation transformants was <2.8 for one-copy LM(ANLGA) plants and <2.1 for one-copy LM9(ANLGA) plants (Table 3). The first-generation transformants originated from tissue culture. Therefore, the samples taken from these plants were not as homogeneous as the samples taken from the genetically identical offspring plants. This difference adds to the variance observed in the populations of first-generation transformants. As a consequence, the 90% upper confidence limit of the variance ratio is likely to be even smaller than the fold difference given above. Therefore, the presence of the MAR sequences had reduced the variation to such an extent that individual, independent one-copy transgenic plants, although characterized by different insertion loci, behaved nearly as if they were a population of genetically identical individuals.

Statistical significance as well as 90% upper confidence limits depend on the sample sizes used. Assuming the observed variances (Table 3), as many as 150 independent, one-copy LM9(ANCGA) transformants are required to reach a 90% upper confidence limit of 1.8, when compared with the same ANCGA-18-He offspring population. Likewise, ~50 independent, one-copy LM(ANLGA) plants have to be generated to reduce the 90% upper confidence level to 2.3. Therefore, further analyses preferably should be performed with the *Lhca3.St.1*–GUS gene. Alternatively, improved statistical methods and/or a gene product that can be measured more accurately than GUS may allow a more powerful determination of variance ratios. Given current technology and sample sizes, however, it can be concluded that the variation in the MAR-containing one-copy plants had approached the lower limit set by random errors in the quantitative GUS measurement. It will be nearly impossible to detect routinely further reductions in variability. Hence, the reduction observed because of the addition of the A elements approaches the maximal possible reduction of GUS transgene variability.

Insertion Site as Hidden Parameter

The eight offspring populations analyzed covered four independent transgene integrations and differed in three characteristics: (1) the presence or absence of the chicken lysozyme MAR element surrounding the T-DNA; (2) the *Lhca3.St.1* promoter or the dCaMV promoter driving the GUS gene; and (3) a homozygous or hemizygous state of the T-DNA. Comparisons with respect to these characteristics indicate the relative role of each character, provided that the plant line chosen can be considered a random, representative example from a hypothetically infinite population of transformants carrying the same T-DNA. The plants that differed in MAR presence (character 1) and/or promoter type (character 2) necessarily also differed in the position of the T-DNA. It could be argued therefore that all observations apply only to these specific integration places. With respect to the populations that differed in homozygous or hemizygous state of the T-DNA (character 3), comparisons obviously involved the same insertion locus.

The question is whether the parent plants chosen could be considered a random sample from our populations containing an intact, single copy of the T-DNA. The limited variation of GUS gene expression that we observed in populations of first-generation MAR-containing plants (Mlynárová et al., 1994, 1995), as well as the results described above, indicated that a randomly chosen individual from such a population could be considered a suitable representative. However, one-copy, non-MAR-containing plants showed considerable variation in GUS gene expression. Therefore, in the case of the one-copy, non-MAR-containing plants (NLG-11-Ho, NLG-11-He, NCG-9-Ho, and NCG-9-He), the specific insertion site also should be taken into account. Assuming that MAR elements shield against the influences of surrounding chromatin, non-MAR-containing populations of plants are likely to be more heterogeneous than the MAR-containing plants. For example, the non-MAR T-DNA may integrate by chance next to potent plant MARs. The resulting transformant is likely to behave as if the T-DNA had MAR sequences. Unfortunately, plants carrying T-DNAs with and without MARs at the same insertion site were not available and could not be obtained readily. To assess the importance of the integration locus for variability, we analyzed the variance in GUS gene activity for eight additional trans-

genic lines without A elements (Table 7). The overall trend indicated that the variance of the NCG-9-He population was reasonably representative. The variability of the NLG-11-He population, however, was the highest of all five NLG populations analyzed (Table 7). The comparisons involving NLG-11-He (Table 6) therefore may suffer from this relatively high variance.

GUS Gene Expression in MAR-Containing versus Non-MAR-Containing Plants

In the case of the *Lhca3.St.1* promoter, the MAR-containing populations (ANLGA-13-Ho and ANLGA-13-He) had the lower activity, whereas in the case of the dCaMV promoter, the MAR-containing populations (ANCGA-18-Ho and ANCGA-18-He) had the higher activity. The high GUS activity of the NLG-11-Ho and NLG-11-He populations came somewhat as a surprise to us but should be attributed to the chance of random position effects. The NLG-11-He population having the highest variance of the five NLG transformants evaluated (Table 7) also indicates chance. The results illustrated that the presence of MARs did not imply higher GUS activity; this finding agrees with our previous observations that the A element per se has no enhancer activity (Mlynárová et al., 1994). This is different from results obtained with the same MAR sequence in animal cell lines (Stief et al., 1989) and also from results obtained with other MAR sequences in plant cell lines (Allen et al., 1993; Spiker and Thompson, 1996). Therefore, either test system (cell line versus mature plants) or particular MAR sequence was important. Analyses of the relative variability in our populations showed that the variation in GUS gene expression levels was lower in plants carrying the MAR elements. Because most variation in these plants was environmental in origin, this demonstrated that the MAR sequences protected against such random variation. It can be concluded that the shielding effect of MARs on the variability of transgene expression observed in the first-generation transformants (Mlynárová et al., 1994, 1995) was transmitted to the next generation.

Gene Expression in *Lhca3.St.1*-Carrying versus dCaMV-Carrying Plants

Unexpectedly, the two large, non-MAR-containing populations carrying the dCaMV promoter–GUS gene T-DNA (NCG-9-Ho and NCG-9-He) had by far the highest variability. Overall, populations carrying the dCaMV promoter–GUS gene tended to be more variable than *Lhca3.St.1* promoter–carrying populations (Figure 2). We had expected the *Lhca3.St.1* promoter to be more sensitive to environmental conditions. This promoter is light regulated (Nap et al., 1993b), whereas the dCaMV promoter is generally considered to be constitutive. This result indicates that the dCaMV promoter itself is intrinsically more susceptible toward variation than is the *Lhca3.St.1* promoter. For example, there may be large differences between individual leaf cells sampled for analysis. The much lower variation in GUS activity observed in the ANCGA-18-He and ANCGA-18-Ho populations (Table 2), as well as in the LM9(ANCGA)-all and LM9(ANCGA)-1copy populations (Table 3), establishes the importance of the MAR sequence on reducing the variability exhibited by the dCaMV promoter. This may indicate that the sensitivity toward variation resides in the nuclear environment of the promoter.

With a simple model of the dynamics of (trans)gene expression, we can postulate that each nucleus establishes upper and lower control limits between which gene expression is allowed to fluctuate. When, for whatever reason, gene expression falls outside of these control limits, it is halted. This regulation is equivalent to the concept of a control chart in industrial quality control (Neter et al., 1988). The high variation observed in the dCaMV promoter–carrying populations (NCG-9-He and NCG-9-Ho) is indicative of a higher susceptibility to transcription-related gene-silencing phenomena. The CaMV promoter has been associated with transgene silencing to such an extent that Matzke and Matzke (1995) suggested that its use be avoided altogether. Inclusion of the dCaMV-directed transgene in a MAR-delimited DNA loop clearly reduced the variation and may therefore decrease the likelihood of silencing.

Gene Expression in Homozygous versus Hemizygous Plants

In all four cases examined, the homozygous plants were approximately twice as active as the corresponding hemizygous plants, showing simple additivity of GUS gene expression. This additivity was independent of the presence of MARs or the promoter type. The additive gene activities indicated that the two allelic copies experienced the same (micro)nuclear environment. Analyses of gene distributions in interphase nuclei suggest that allelic copies of a gene occur in different parts of the nucleus (Heslop-Harrison, 1992). The additive gene activity indicates that nuclei are divided into subcompartments that create identical conditions for the individual alleles to be expressed. The chromosome structure around the allele may be responsible for such conditions. Part of the higher order chromatin regulation of gene expression may be made up of a mechanism that duplicates nuclear conditions at different sites of the nucleus.

Previously, we demonstrated that first-generation transformants carrying the pLM(ANLGA) T-DNA (MAR–*Lhca3.St.1*-GUS gene) show a copy number dependence of GUS gene expression (Mlynárová et al., 1994). In contrast, first-generation transformants carrying the pLM9(ANCGA) construct (MAR–dCaMV–GUS gene) show some form of dosage compensation: transformants have the same approximate GUS activity, irrespective of the number of gene copies integrated (Mlynárová et al., 1995). The results obtained here with the ANCGA-18 and NCG-9 populations appear to contradict the existence of such a dosage compensation. The apparent contradiction may indicate that allelic and ectopic copies of the transgene behave differently. However, in independent,

homozygous, MAR-containing and dCaMV–GUS gene–carrying transgenic plants, we have observed a gradual decrease of GUS gene activity to approximately the level of the corresponding hemizygous plants (Ľ. Mlynárová, unpublished observations). The supposed dosage compensation mechanism may therefore be related to either the age of the plant or any stress related to aging.

Gene Variability in Homozygous versus Hemizygous Plants

For the MAR-containing plants (ANLGA-13 and ANCGA-18), the homozygous populations had approximately a twofold higher variation in GUS gene activity than did the corresponding hemizygous populations (Table 6). The variance of a sum equals the sum of the variances only when the variables are uncorrelated (Sokal and Rohlf, 1995). The approximately twofold higher variance observed in the homozygous populations therefore indicates that the MAR-contained GUS gene alleles not only are additive but also act fully independent from each other. The "control chart" model outlined above predicts that the higher relative variabilities observed in homozygous populations will result in a higher susceptibility of homozygotes toward gene silencing. The MAR-mediated protection against gene silencing by reducing variation will have to compete with increased variability in the homozygous state. Indeed, homozygous (transgenic) plants in particular are found to be susceptible to gene-silencing phenomena (Hart et al., 1992; Dorlhac de Borne et al., 1994; Brandle et al., 1995).

METHODS

T-DNA Vectors and Plant Transformation

Transgenic tobacco (*Nicotiana tabacum* cv Petit Havana SR1) plants, obtained after *Agrobacterium tumefaciens*–mediated transformation with the T-DNA vectors pPPG, pLM, pLM5(NG), and pLM9(ANGA), originated from previous studies (Mlynárová et al., 1994, 1995). All T-DNA vectors are derivatives of pBin19 (Bevan, 1984), carrying a kanamycin resistance gene encoding aminoglycoside phosphotransferase A2 (*aph*A2) (synonymous with neomycin phosphotransferase II [*npt*II]) for selection. In addition, they carry either a *Lhca3.St.1* promoter–β-glucuronidase (GUS) gene fusion (pPPG and pLM) or a doubled cauliflower mosaic virus 35S (dCaMV) promoter–GUS gene fusion (pLM5[NG] and pLM9[ANGA]); with chicken lysozyme A elements at each T-DNA border (pLM and pLM9[ANGA]), or without these A elements (pPPG and pLM5[NG]).

Nomenclature of Plants and Populations

The characteristics of the plant material used in this study are summarized in Table 1. The parent plants selected for the large-scale offspring analyses cover four T-DNA configurations, of which populations of primary transformants have been analyzed in detail (Mlynárová et al., 1994, 1995). To facilitate identification throughout the text, the plant lines involved were named after the T-DNA configuration they carried. Plants carrying the pLM T-DNA were designated ANLGA, plants with the pPPG T-DNA were designated NLG, plants carrying pLM5(NG) were designated NCG, and plants carrying pLM9(ANGA) were designated ANCGA. In this system, A represents the A element, N the kanamycin resistance gene, L the *Lhca3.St.1* promoter (Nap et al., 1993b), C the dCaMV promoter, and G the GUS gene. The plant lines chosen for analysis of large progeny populations comprised ANLGA-13, NLG-11, ANCGA-18, and NCG-9. All four parent plants carried a single, intact integration of the T-DNA, according to DNA gel blot and genetic segregation analyses (data not shown).

Homozygous plants were identified among the progeny of the primary transformants on the basis of absence of segregation for kanamycin resistance after self-pollination and backcrosses to wild-type untransformed tobacco. The homozygous plant lines were maintained by self-pollination. The lines used were designated NLG-11-Ho, ANLGA-13-Ho, NCG-9-Ho, and ANCGA-18-Ho. Homogeneous seedlots of hemizygous plants were obtained by backcrossing the homozygous plants to the wild type. The hemizygous plant populations obtained were designated NLG-11-He, ANLGA-13-He, NCG-9-He, and ANCGA-18-He, respectively. More than 100 seeds were sown from all eight populations. Three weeks after germination, ~45 randomly selected seedlings were transferred to individual pots. Plants were assayed for GUS activity after an additional 2 weeks of growth. To minimize variation due to environment, plants were treated as uniformly as possible and grown in a fully climatized greenhouse with heating below 18°C and ventilation above 24°C.

To assess the importance of the integration locus, we chose four additional plant lines for the non–matrix-associated region (MAR) T-DNAs, NLG and NCG. All lines carried a single, intact integration of the T-DNA according to DNA gel blot and genetic segregation analyses (data not shown). The plant lines chosen were indicated NLG-1, NLG-10, NLG-15, NLG-41, NCG-4, NCG-10, NCG-47, and NCG-52 (Table 1). Over 50 seeds were sown from the selfed primary transformants. Thirty plants were transferred to individual pots and assayed for GUS activity after an additional 3 weeks of growth. Hemizygous plant lines were identified on the basis of a 3:1 segregation for kanamycin resistance after self-pollination.

The characteristics of the populations of the A element–carrying primary transformants previously described (Mlynárová et al., 1994, 1995) are also given in Table 1. To be consistent with the system of plant nomenclature described above, these populations were designated LM(ANLGA) when they were carrying the T-DNA from pLM (Mlynárová et al., 1994) and LM9(ANCGA) when they were carrying the pLM9(ANGA) T-DNA (Mlynárová et al., 1995). The suffix "-all" gives the full population of first-generation transgenic plants obtained; the suffix "-1copy" indicates the subset of plants carrying one copy of the T-DNA.

GUS Assays and Statistical Analysis

GUS measurements were performed as described by Mlynárová et al. (1994), using a Fluoroskan II microtiterplate reader (Titertek, Finland). Samples were harvested as 9-mm-diameter discs cut through similar positions of the lamina from leaves of the same age and position. All statistical calculations were performed with the program Genstat 5 (Payne et al., 1993) and interpreted using the statistical tables of Rohlf and Sokal (1994).

ACKNOWLEDGMENTS

We thank the CPRO–DLO greenhouse team for excellent care of the plants; Dr. Tony Conner (OECD collaborator, Christchurch, New Zealand) for making the relevant crosses and for discussions and invaluable help with the manuscript; Drs. Fred van Eeuwijk, Ritsert Jansen, and Peter Metz for discussions and support; and Drs. Andy Pereira, Ruud de Maagd, and Dirk Bosch for commenting on the manuscript. The investigations were supported by the Life Sciences Foundation (SLW), which is subsidized by the Netherlands Organization for Scientific Research (NWO) and by program subsidy 280 from the Dutch Ministry of Agriculture. Ľ.M. was on leave from the Institute of Plant Genetics, Nitra, Slovak Republic.

Received April 5, 1996; accepted July 8, 1996.

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.

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

Brandle, J.E., McHugh, S.G., James, L., Labbé, H., and Miki, B.L. (1995). Instability of transgene expression in field grown tobacco carrying the *csr1-1* gene for sulfonylurea herbicide resistance. Bio/Technology. **13,** 994–998.

Conner, A.J., and Christey, M.C. (1994). Plant breeding and seed marketing options for the introduction of transgenic insect-resistant crops. Biocontrol Sci. Technol. **4,** 463–473.

Conover, W.J. (1980). Practical Nonparametric Statistics, 2nd ed. (New York: John Wiley and Sons).

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

Dorlhac de Borne, F., Vincentz, M., Chupeau, Y., and Vaucheret, H. (1994). Cosuppression of nitrate reductase host genes and transgenes in transgenic tobacco plants. Mol. Gen. Genet. **243,** 613–621.

Falconer, D.S. (1981). Introduction to Quantitative Genetics, 2nd ed. (London: Longman).

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.

Hart, C.M., Fischer, B.M., Heuhaus, J., and Meins, F., Jr. (1992). Regulated inactivation of homologous gene expression in transgenic *Nicotiana sylvestris* plants containing a defense-related tobacco chitinase gene. Mol. Gen. Genet. **235,** 179–188.

Heslop-Harrison, J.S. (1992). Nuclear architecture in plants. Curr. Opin. Genet. Dev. **2,** 913–917.

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.

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.

Lane, P.W., and Payne, R.W. (1994). An Introductory Course for Genstat 5 Release 3. (Rothamsted, UK: Lawes Agricultural Trust–Rothamsted Experimental Station).

Lehmann, E.L. (1986). Testing Statistical Hypotheses, 2nd ed. (New York: John Wiley and Sons).

Matzke, M.A., and Matzke, A.J.M. (1993). Genomic imprinting in plants: Parental effects and *trans*-inactivation phenomena. Annu. Rev. Plant Physiol. Plant Mol. Biol. **44,** 53–76.

Matzke, M.A., and Matzke, A.J.M. (1995). How and why do plants inactivate homologous (trans)genes? Plant Physiol. **107,** 679–685.

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.

Mlynárová, Ľ., Jansen, R.C., Conner, A.J., Stiekema, W.J., and Nap, J.-P. (1995). The MAR-mediated reduction in position effect can be uncoupled from copy number–dependent expression in transgenic plants. Plant Cell **7,** 599–609.

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 *Lhca.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.

Neter, J., Wasserman, W., and Whitmore, G.A. (1988). Applied Statistics, 3rd ed. (Boston: Allyn and Bacon, Inc.).

Payne, R.W., Lane, P.W., Digby, P.G.N., Harding, S.A., Leech, P.K., Morgan, G.W., Todd, A.D., Thompson, R., Tunnicliffe Wilson, G., Welham, S.J., and White, R.P. (1993). Genstat 5 (Release 3) Reference Manual. (Oxford, UK: Clarendon Press).

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.

Rohlf, F.J., and Sokal, R.R. (1994). Statistical Tables, 3rd ed. (New York: Freeman and Co.).

Sokal, R.R., and Rohlf, F.J. (1995). Biometry, 3rd ed. (New York: Freeman and Co.).

Spiker, S., and Thompson, W.F. (1996). Nuclear matrix attachment regions and transgene expression in plants. Plant Physiol. **110,** 15–21.

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