Plant Molecular Biology **15**: 851–864, 1990.

The effect of T-DNA copy number, position and methylation on reporter gene expression in tobacco transformants

Shaun L. A. Hobbs, Pascal Kpodar[1] and Catherine M. O. DeLong
Plant Biotechnology Institute, National Research Council of Canada, 110 Gymnasium Road, Saskatoon, Saskatchewan, S7N OW9, Canada; [1]Present address: Centre de Physiologie Végétale, Unité Associée au C.N.R.S. no. 241, 118 route de Narbonne, 31062 Toulouse Cedex, France

Received 9 May 1990; revised and accepted 30 August 1990

Key words: CaMV-35S promoter, *Agrobacterium tumefaciens, Nicotiana tabacum* L., GUS activity, inter-transformant variability

Abstract

Inter-transformant variability in the expression of introduced genes was studied in the R_1 and R_2 generations of 10 tobacco transformants, produced by *Agrobacterium*-mediated transformation. In replicated and physiologically equivalent material, tranformants showed considerable variability in the expression of the reporter gene *uidA* as shown by transcript levels and β-glucuronidase (GUS) activity. However, homozygous R_2 material could be investigated for seven of the transformants and among these, and in one line in which two inserts could segregate independently, this inter-transformant variability was reduced to simple bimodal expression. The two levels of expression for GUS activity in leaves were high or low (approximately 2.5 or 0.3 nmol cm^{-2} min^{-1} respectively), with no continuous variation. Transformants in the high group had single T-DNA insertions, while those in the low group had multiple T-DNA insertions, at the same or different loci. Within each group, although T-DNA was apparently integrated at different sites in the plant genome, there was no evidence of position effects. GUS activity levels of the transformants were very similar in the field and in environmentally controlled conditions under high or low light. Plants with multiple insertions and low expression also tended to have increased methylation of the integrated T-DNA.

Introduction

Plant transformants produced using *Agrobacterium*-mediated techniques can vary widely in the expression of the introduced genes. The level of protein coded for by an alien gene commonly varies 10 to 50 fold among individual transformants within the same experiment [3, 15, 19, 20, 24, 27, 29, 30] and substantial inter-transformant variability also occurs in the amount of gene transcript [4, 8, 15, 19, 20, 24, 35].

Such inter-transformant variability has been suggested to be due to a variety of genetic causes, none of which fully explains the phenomenon. Copy number of the introduced T-DNA varies among transformants but it has often been shown that there is no positive correlation between increased copies and increasing expression of the genes in the T-DNA [4, 15, 16, 19, 27, 29, 30] and there are even indications that single copies result in higher expression [16, 27, 29]. The insertion of T-DNA is random within the plant genome and

the activity of the introduced genes may be affected by adjacent plant DNA (position effect); however, this has been found not to be the simple, exclusive cause of variable expression in many instances [4, 8, 16]. Truncation, rearrangement or repetition of the introduced T-DNA may also affect gene expression but there has not always been a direct correlation shown between low expression and deletions, rearrangements [7], or inverted repeats [16].

Inter-transformant variability can be caused by experimental or biological 'error', as well as by defined genetic variability, and these errors are not always fully controlled in inter-transformant comparisons. First, epigenetic effects can cause substantial variability in the primary regenerants (R_0), with up to 16 fold differences in gene expression being reported among genotypically identical clones [16]. Second, spatial and temporal differences in gene expression can occur if plant material is not physiologically equivalent [3, 4, 16, 35]. Third, as in any biological assay, the effect of environment and plant-to-plant variability necessitates that replicated material should be used along with multiple assays of quantitative characters so that suitable statistics can be obtained. Such effects are especially important if the promoter used to control the expression of the introduced gene is affected by the environment (as reviewed by Benfey and Chua [1]) or if the assay for the gene product is particularly imprecise. Fourth, transformation can produce the insertion of single or multiple copies of full-length, truncated or rearranged T-DNA, in either orientation, at one or more loci in the plant genome [7, 16, 17, 32]. Therefore, to avoid confounding one putative genetic effect with another, investigations of position effects should compare only plants of similar composition but with different locations of insertions.

Few, if any, of the investigations to date involving inter-transformant variability in gene expression have taken all these factors into account. In the work presented here, *Nicotiana tabacum* L. was transformed using pBI121 [14] and inter- and intra-transformant variability of β-glucuronidase (GUS) activity was compared in replicated, physiologically synchronised R_1 and R_2 plants in three different environments, including the field. The data indicated that, among homozygous material, inter-transformant variability consisted only of high or low GUS activity, with no continuous variation. Those plants with single T-DNA insertions had high expression whereas those plants with multiple insertions, at one or more loci, had low expression which was associated with increased methylation of the integrated T-DNA. There was no indication of position effects within the high or low group.

Materials and methods

Plant transformation

The leaf disc transformation method [12] was used. Briefly, 7 mm diameter leaf discs from surface-sterilised leaves of *Nicotina tabacum* L. cv. Xanthi were cocultivated with disarmed *Agrobacterium tumefaciens* strain C58 [18] containing the binary vector pBI121 [14]. The T-DNA of pBI121 contains *kan* (coding for neomycin phosphotransferase II) driven by a *nos* promoter and *uidA* (coding for β-glucuronidase (GUS)) driven by a cauliflower mosaic virus (CaMV)-35S promoter (Fig. 1A). After 2 days on cocultivation medium of MS salts [23], B5 vitamins [9], 2 mg l^{-1} 2,4-dichlorophenoxyacetic acid (2,4-D), 3% sucrose, 0.8% agar, pH 5.6 the discs were removed to regenerating, selective medium (as above but with 2,4-D replaced by 2.5 mg l^{-1} benzyladenine (BA) and 0.1 mg l^{-1} α-naphthalene acetic acid (α-NAA) and supplemented with 500 mg l^{-1} carbenicillin and 100 mg l^{-1} kanamycin sulphate). Small (approximately 1.5 cm) plantlets were removed as they appeared and were placed on rooting medium (as above but with only 0.1 mg l^{-1} α-NAA as hormone), and were transferred to soil after roots appeared. Pieces of leaves from transformants were placed on selective callusing medium (MS salts, B5 vitamins, 2.0 mg l^{-1} 2,4-D, 0.5 mg l^{-1} BA, 100 mg l^{-1} kanamycin sulphate) and any escapes, i.e. those not forming callus, were discarded.

Table 1. GUS activity (nmol per mg protein per min) among 20 R_1 plants from 10 transformants.

Transformant number	Mean[a]	Range	Highest/lowest[b]	Mean of homozygous lines[c]	Mean of hemizogous lines[d]
T3	6.7 ± 1.2	1.1–20.6	19		
T4	1.6 ± 0.1	0.5–3.3	6	1.3 ± 0.3 L	1.7 ± 0.3
T5	5.6 ± 0.5	2.9–9.6	3	8.2 ± 0.7 H	4.4 ± 0.5
T6	1.8 ± 0.4	0.1–7.9	80		
T7	3.9 ± 0.5	0.7–6.5	9	2.4 ± 1.4 L	4.9 ± 0.5
T11	0	–	–		
T13	5.5 ± 0.6	3.2–11.6	4	8.6 ± 0.6 H	4.4 ± 0.2
T14	5.2 ± 0.5	34–8.1	2	6.4 ± 0.4 H	4.1 ± 0.2
T18	1.3 ± 0.1	0.7–1.9	3	1.1 ± 0.2 L	1.3 ± 0.2
T19	6.6 ± 0.6	3.5–12.8	4	7.7 ± 0.8 H	4.7 ± 0.3

[a] ± standard error. Means of all R_1 plants except those with no GUS activity.
[b] GUS activity of highest R_1 plant as a ratio of that of lowest.
[c] ± standard error. Means of homozygous GUS-positive R_1 plants only. H and L indicate homozygous transformants which appear to have high or low GUS expression, respectively.
[d] ± standard error. Means of hemizygous GUS-positive R_1 plants only.

R_1 progeny

Nomenclature of transformed material is as in Potrykus [27]. Ten primary (R_0) transformants (identified in Table 1) were selected and 30 R_1 seeds from each, bulked from several different seed capsules collected from bagged heads, were surface-sterilised (30 s treatment in 10% commercial bleach (0.6% sodium hypochlorite)) and rinsed well in sterile water. Seeds were then sown onto solid selective germination medium (1/2 strength MS salts, B5 vitamins and 100 mg l^{-1} kanamycin sulphate) and seedlings scored for susceptibility (bleached) and resistance (green) to kanamycin after 28 days. In our hands, this level of kanamycin gave good separation into resistant and susceptible types. This qualitative test was the only test used for kanamycin resistance, there being no attempt to quantify *kan* expression levels. R_1 seed was also sown in 20 pots (20 cm) for each transformant in Terra-Lite Redi-Earth (W.R. Grace and Co., Canada) and thinned to one plant per pot. Plants were grown under 250 μmol quanta m^{-2} s^{-1} photosynthetically active radiation (PAR) with a 16 h photoperiod and 22 °C/18 °C day/night ambient temperature. Ninety days after germination, single leaf discs (8 mm diameter) were punched from the middle of the lamina of the newest fully expanded leaf, avoiding major veins, and GUS activity was measured using 4-methylumbelliferyl-β-D-glucuronide (MUG) as substrate and methylumbelliferone as standard [13] and protein levels were determined using a Bio-Rad (Richmond, CA, USA) protein assay kit. Leaf discs were always taken from the same place on leaves of a uniform size sampled at the same position on the plant. R_2 seed was collected from several surface-sterilised (wiped with 70% ethanol) seed capsules from bagged heads and kept separately for each plant. In this way 20 R_2 populations (lines), numbered L1 to L20 after each of the 20 R_1 plants, were generated for each transformant.

R_2 progeny

Approximately 100 seeds from each of the 200 R_2 lines were germinated on selective medium and scored as above. Seed from 26 lines, with representation from each individual transformant, was taken for further investigation (Table 2). Selected lines included those that: were from transformants where a single locus had been affected and could be discerned to be homozygous for introduced T-DNA (T4, T5, T7, T11, T13, T14, T18, T19); were expected to be segregating for multiple

Table 2. GUS activity in leaves of field-grown R_2 lines.

Transformant	Line no.	GUS cm^{-2} min^{-1}	
		Mean[a]	Coefficient of variation
T3	3[b]	0.81 ± 0.24	75
T3	13[b]	0.83 ± 0.31	91
T3	18[b]	2.61 ± 0.18	17
T4	5	0.33 ± 0.05 L	36
T4	12	0.35 ± 0.05 L	36
T4	17	0.36 ± 0.05 L	33
T5	5	2.40 ± 0.14 H	15
T5	9	2.24 ± 0.13 H	14
T5	18	2.48 ± 0.13 H	13
T6	3[b]	0.22 ± 0.07	78
T6	8[b]	0.59 ± 0.19	78
T6	18[b]	0.04 ± 0.02	100
T7	6	0.45 ± 0.06 L	26
T11	15	0.00	
T11	20	0.00	
T13	15	2.42 ± 0.17 H	17
T13	17	2.29 ± 0.07 H	8
T14	2	2.40 ± 0.14 H	15
T14	19	2.52 ± 0.14 H	13
T18	11	0.34 ± 0.05 L	36
T18	19	0.30 ± 0.02 L	16
T19	10	2.58 ± 0.08 H	8
T19	11	2.52 ± 0.08 H	8
T19	13	2.48 ± 0.13 H	13
T4	10[c]	0.00	
T19	20[c]	0.00	

[a] ± standard error. H and L indicate homozygous lines which appear to have high or low GUS expression, respectively.
[b] Line was not necessarily homozygous.
[c] Control line.

copies of T-DNA at different loci (T3 and T6) and; were expected to have had their T-DNA segregated out (T4: L10 and T19: L20). Seed from each selected line was germinated in 36 pots (5 cm squared), thinned to one seedling per pot and grown under the above environmental conditions. Forty-two days after sowing, plants of uniform size across the selected lines were chosen and transplanted as follows: one plant of each to 6 replicate blocks in the field; one plant of each to 6 replicate blocks under the above controlled environment conditions; and one plant of each to 4 replicate blocks under the above controlled environment conditions except with a PAR of 800 μmol quanta $m^{-2}s^{-1}$. Lines were randomised within each block. Permission for field testing of transgenic material was obtained from the Plant Health Directorate of the Seed Division of Agriculture Canada, the regulatory body for such tests in Canada.

Between 14 and 28 days after transplanting, leaves from all plants in all environments were assayed twice (approximately at the 5–6 and 7–8 leaf stages) for GUS activity as above. In one replicate block in the field, those leaves that were sampled for activity were also harvested, ground in liquid nitrogen and stored at − 70 °C for later DNA and RNA analysis. Leaves were similarly harvested for all 16 plants (all blocks and all environments) of T3 : L3.

Southern and northern hybridisations

DNA was extracted [6] from representative plants from each transformant, digested with restriction enzymes according to the manufacturer's recommendations and Southern blots and hybridisations performed [11]. The ^{32}P-labelled probes used from pBI121 were the 1.9 kb (kilobase) *Bam*HI-Sst I fragment and the 0.4 kb *Pst*I-*Sph*I fragment for *uidA* and *kan* respectively (Fig. 1A). Extractions and blots were repeated at least once.

RNA was extracted [21] from the same leaf material and approximately 10 μg of total RNA was used to perform northern blots, and 10, 5, 2 and 1 μg were used for dot blotting on GeneScreen Plus (NEN Research Products, Boston, MA, USA) according to the instruction booklet.

Statistical analysis

Coefficients of variation were calculated and analyses of variance performed according to Snedecor and Cochran [31].

Results

R_1 progeny assays

To avoid epigenetic effects on primary transformants, R_1 and R_2 generations were examined. Twenty R_1 plants from each of 10 randomly selected kanamycin-resistant tobacco transformants were grown under controlled environment conditions. The expression of *uidA* in the leaves of these transformants varied considerably as shown by GUS activity levels (Table 1) with one, T11, showing no GUS activity. The difference between the highest (20.6 nmol per mg protein per min) and the lowest (0.1 nmol per mg protein per min) positive R_1 plants was greater than 200 fold. Repeat assays for low, positive T6 plants showed four differing from 0 but with activities below 0.3 nmol per mg protein per min. Within each transformant, GUS activity of the individual R_1 plants differed but the ratio of highest to lowest plant (excluding those with zero GUS activity) was greater for T3 and T6 than for the others (Table 1). Chi-square tests on results of sensitivity to kanamycin were consistent with the hypothesis that all transformants except T3 and T6 had been affected at a single locus. The data for T3 fitted a model for insertion at 2 loci whereas that for T6 did not fit either simple model.

Testing R_2 seedlings for kanamycin resistance determined the genetic composition of each individual R_1 plant, i.e. whether homozygous for inserted T-DNA (R_2 all resistant), homozygous for no T-DNA (R_2 all susceptible), or hemizygous (R_2 segregating resistant and susceptible). For each transformant with T-DNA insertion(s) at a single locus, the previously measured positive R_1 GUS activities were split into means for hemizygous and homozygous plants. Means for homozygous plants generally appeared to agree with overall means (Table 1) in that 2 groups were formed, one with high GUS activity (T5, T13, T14 and T19) and one with low activity (T4, T7 and T18). In the high group, the activity of the hemizygous plants averaged approximately half that of the homozygous plants (Table 1). In the low group there was no significant difference between hemizygous and homozygous averages.

R_2 progeny assays under different environmental conditions

Seed from a maximum of 3 homozygous T-DNA positive R_1 plants was chosen for each transformant to produce replicated homozygous R_2 lines for assay under different environments. Lines from T3 and T6 were also included, even though these were not expected to be homozygous, as were GUS-negative control lines (T4 : L10 and T19: L20). Plants were grown under controlled conditions of low light and high light as well as in the field. Differences among transformants were very stable across environments. Not only were rankings of transformants very similar but so were the absolute GUS activity levels (per unit leaf area) for each transformant (Fig. 2). Variability within each transformant was higher for T3 and T6 as shown by the larger coefficient of variation (Fig. 2). GUS activity expressed on a

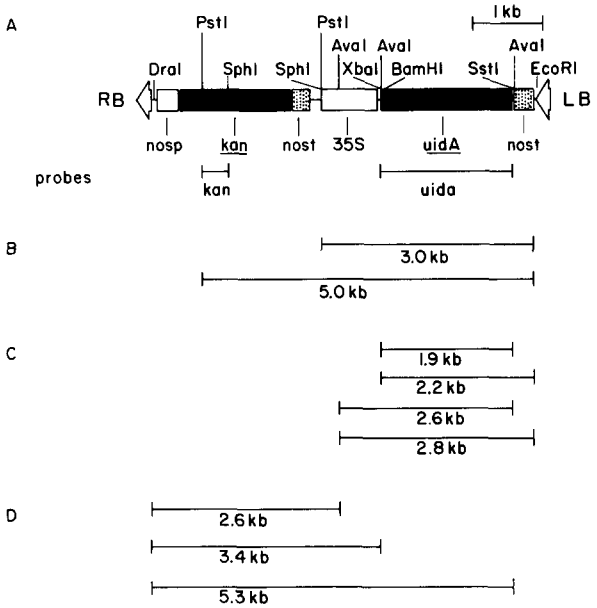

Fig. 1. T-DNA from pBI121. A, Arrangement of chimaeric genes, restriction sites and probes used (nosp = *nos* promoter, nost = *nos* terminator, 35S = CaMV-35S promoter, RB = right border, LB = left border). B to D, Possible restriction fragments produced within the T-DNA by digestion with: *Eco*RI together with *Pst*I when probed with *uidA* (B); *Eco*RI together with *Ava*I when probed with *uidA* (C); and *Dra*I together with *Ava*I when probed with *kan* (D).

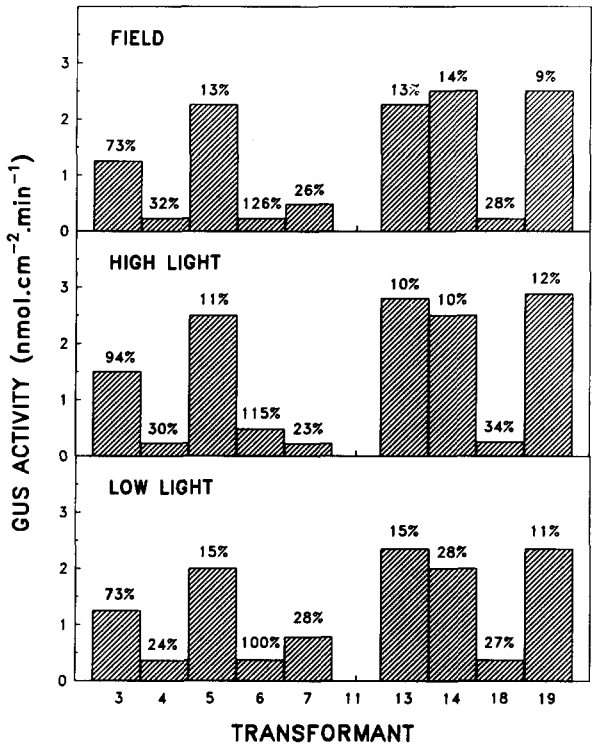

Fig. 2. Mean GUS activity in R_2 plant leaves from 10 transformants in different environments. Figures above columns are coefficients of variation.

per mg protein basis also showed transformant rankings to be similar within each environment but absolute values differed among environments (mean of homozygous lines = 3.2, 5.4 and 5.3 nmol per mg protein per min in the field, high light and low light environment respectively). These environmental differences were caused by differing overall leaf protein levels (0.5, 0.3 and 0.3 mg protein cm^{-2}, respectively.

The GUS activities for individual lines from each transformant grown in the field are given in Table 2, data from other environments were very similar (data not presented). As with R_1 plants, when positive homozygous material was compared, a group with high activity (all lines within T5, T13, T14 and T19) and a group with low activity (all lines within T4, T7 and T18) was found, with no continuous variability. Lines from T11 and controls showed no GUS activity. The coefficients of variation show that variability within each homozygous line was generally considerably lower than that in T3 and T6 lines. This confirmed that lines within T3 and T6 (with the possible exception of T3 : L18) were not homozygous and were still segregating.

Analysis of T-DNA in homozygous transformants

The T-DNA of pBI121 has unique restriction sites for *Eco*RI and *Bam*HI, respectively just inside the left border and between CaMV-35S and *uidA* (Fig. 1A). DNA from leaves measured for GUS activity was extracted from representative R_2 plants from each transformant growing in the field, cut with *Eco*RI and *Bam*HI and Southern blots were probed with *uidA*. All transformants had a complete *Bam*HI-*Eco*RI fragment of approximately 2.1 kb (Fig. 3A) except those with no GUS activity (T11 and T4 : L10).

The number of T-DNA insertions in an individual transformant was determined by DNA restriction mapping using *Eco*RI or *Dra*I. These have unique restriction sites at either end of the T-DNA (Fig. 1A) and hence border fragments between T-DNA and plant DNA should be generated after digestion with each and probing with either *uidA* or *kan*. Insertions were therefore shown as individual bands and thus could be counted. Use of both restriction enzymes independently ensured that any truncation, rearrangement or replication of T-DNA could be detected and would not confuse copy number estimation.

For those lines known to be homozygous (and involving insertions at a single locus), the group showing high GUS activity all had single T-DNA insertions, whereas the group with low GUS activity showed indications of two copies of T-DNA with one or both of the restriction enzymes (Fig. 3B and 3C). The highest molecular weight band in the *Eco*RI digest of T18 : L11 DNA was caused by undigested DNA in this particular blot and was not present in repeat blots. Each plant, except T11, had at least one full-length copy of the T-DNA as determined by digestion with both enzymes at the same time (data not presented). T11 was shown to have a single, partly deleted T-DNA insert, as blots

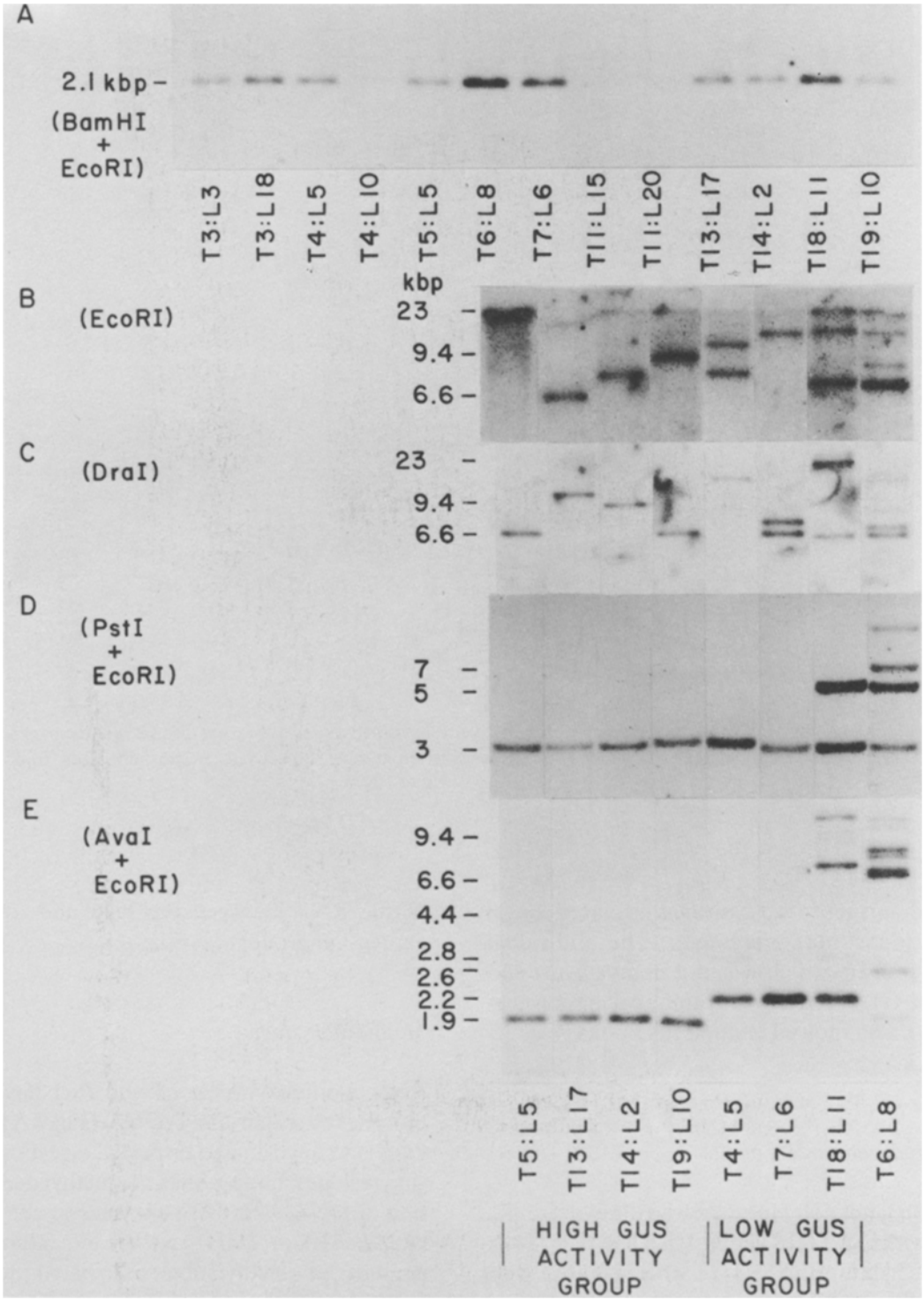

Fig. 3. Autoradiograms of Southern blots of DNA extracted from representative transformed plants (all except T6 : L8, T3 : L3 and T3 : L18 are homozygous), digested with restriction enzymes and filters hybridised with *uidA* probe. A to E, DNA digested with *Eco*RI together with *Bam*HI (A); *Eco*RI (B); *Dra*I (C); *Eco*RI together with *Pst*I (D); and *Eco*RI together with *Ava*I (E).

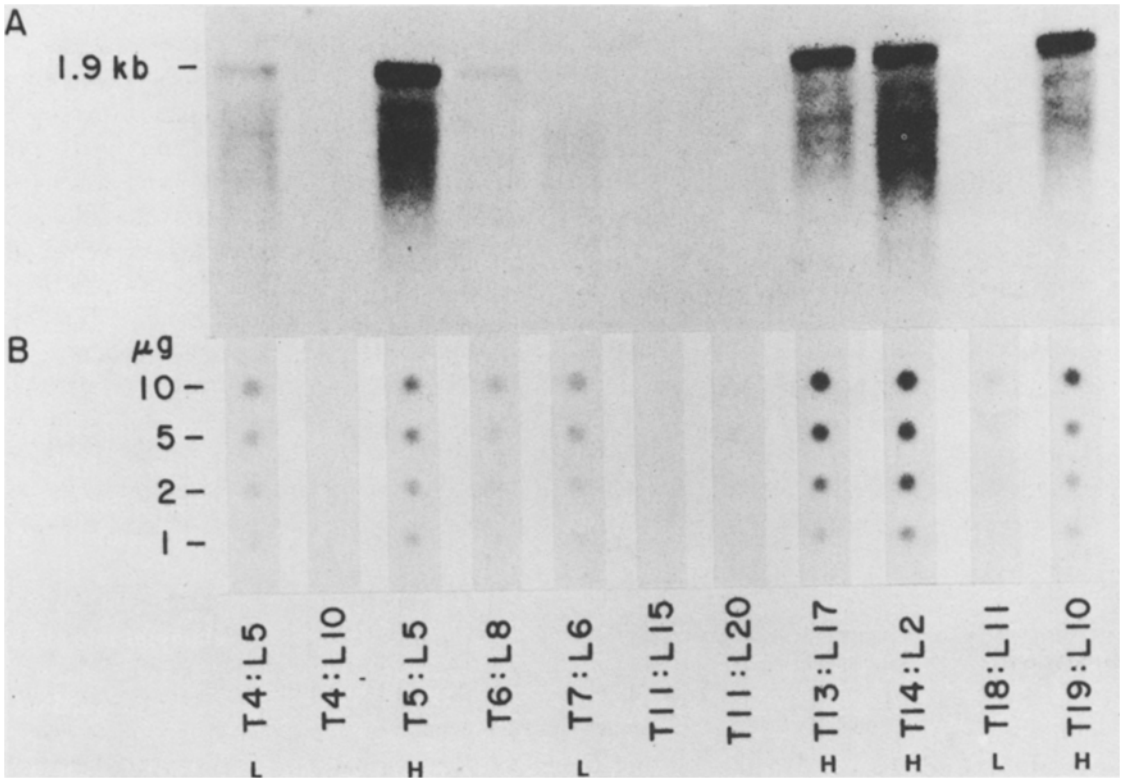

Fig. 4. Autoradiograms of RNA extracted from representative transformed plants (all except T6 : L8 are homozygous) probed with *uidA*. H and L indicate plants from high or low GUS expression groups, respectively. A: northern blots. B: dot blots for 10, 5, 2 and 1 μg of RNA.

stripped and reprobed with *kan* produced single bands (data not presented). The individual example of T6, which was not homozygous and had low GUS activity (0.8 nmol per mg protein per min), also showed multiple insertions (Fig. 3B and 3C).

RNA production in R_2 plants

Northern and dot blots showed that those R_2 plants with high GUS activity also had high levels of RNA hybridisable to *uidA* whereas those with low GUS activity had low levels of hybridisable RNA (Fig. 4A and 4B). The control and T11 had no *uidA* signal. The northern blots (Fig. 4A) also show that there is no obvious difference in the size of the RNA between the high and low GUS expression groups.

Methylation studies

Restriction enzymes *Ava*I and *Pst*I have recognition sites within the T-DNA (Fig. 1A) but are known to have limited or partial digestion if cytosine residues in these sites are methylated. Digestion of DNA with these enzymes in conjunction with *Eco*RI or *Dra*I gave an indication of the amount of methylation associated with the T-DNA.

When genomic DNA from a transformed plant was digested with *Eco*RI and *Pst*I together, a 3.0 kb fragment would hybridise to *uidA* if the

*Pst*I site just 3′ to the CaMV-35S promoter were not methylated (Fig. 1B). If this site were partially or totally methylated, and the *Pst*I site in the *kan* gene were not, a 5.0 kb fragment would be produced (Fig. 1B). If this latter site were also methylated, or partially so, higher molecular weight fragments would be generated. The 3.0 kb fragment was present in all homozygous plants (Fig. 3D). However, the 5.0 kb band was also present in T18 indicating partial or total methylation of the *Pst*I site just 5′ to the CaMV-35S promoter in at least one of its T-DNA copies (Fig. 3D). Methylation was also increased in T6 (Fig. 3D).

Expected fragment sizes following digestion with both *Ava*I and *Eco*RI and probing with *uidA* are shown in Fig. 1C. All homozygous transformants from the high GUS activity group (T5, T13, T14 and T19) showed no indication of methylation, producing only 1.9 kb bands (Fig. 3E). Those homozygous transformants from the low GUS activity group were all methylated at the *Ava*I site between *uidA* and its *nos* terminator (producing 2.2 kb bands). In addition, T4 was partially methylated at the *Ava*I site at the beginning of *uidA* (producing a 2.8 kb band) and T18 was methylated or partially methylated at all *Ava*I sites in the T-DNA, producing some fragments that were the same size as those produced by digestion with *Eco*RI alone (compare Figs. 3B and 3E). Similarly, T6 showed methylation of T-DNA (Fig. 3E).

A segregating R_2 line

As there was considerable similarity between GUS activities per unit leaf area among environments (Fig. 2) it could be assumed that the activity of an individual plant was mainly due to its genetic composition rather than to any environmental effect. Therefore, GUS activities of all 16 individual R_2 plants (all blocks and all environments) within each of the T3 lines were analysed together. One line, T3 : L3, showed a large range (from 0.23 to 2.55 nmol cm^{-2} min^{-1}) with a statistically separable high and low group of plants (respectively greater than 1.91 and less than 0.86 nmol cm^{-2} min^{-1}) (Fig. 5). Copy number and methylation investigations were carried out as before on all the T3 : L3 plants and again each plant with high activity was shown to have a single T-DNA insertion whereas each plant with low activity had an extra T-DNA copy (Figs. 5A and 5B). The exception to this was the appearance of an aberrant low molecular weight band in plant 3 from the low light environment. This fragment was not full-length, hybridising to *uidA* but not to *kan* probes and giving a band of 3.8 kb (in addition to a full-length 5.5 kb fragment) in blots of DNA digested with *Dra*I and *Eco*RI together (data not presented). This fragment also did not appear in the parental (R_1) DNA (Figs. 5A and 5B).

A chi-square test of the T-DNA copy number in the R_2 plants was consistent with the hypothesis that there were two independent copies of T-DNA in T3 : L3, the R_1 parent giving rise to this line being homozygous for T-DNA at locus 1 and hemizygous at locus 2 (corresponding to the lower and higher molecular weight fragments respectively with *Eco*RI or *Dra*I) hence producing an expected 3 : 1 ratio of R_2 plants with T-DNA at both loci to those with T-DNA only at locus 1. Presence of T-DNA at both loci produced low GUS activity, presence of T-DNA only at locus 1 produced high activity. The level of GUS activity associated with presence of T-DNA at locus 2 and not locus 1 could not be determined as none of the plants in any of the T3 lines examined were found to have this combination.

Digestions of DNA involving either *Ava*I or *Pst*I in conjunction with *Eco*RI showed that plants with high GUS activity generally had less methylation at *Ava*I and *Pst*I sites than low expressors (Figs. 5C and 5D). However, none showed a pattern indicative of total absence of methylation at all sites examined, as no single 1.9 kb bands were produced in the digestions using *Ava*I together with *Eco*RI.

The possible fragment sizes following digestion of DNA with *Ava*I and *Dra*I together and probing with *kan* are shown in Fig. 1D. All 5 of the T3 : L3 plants with high GUS activity showed no

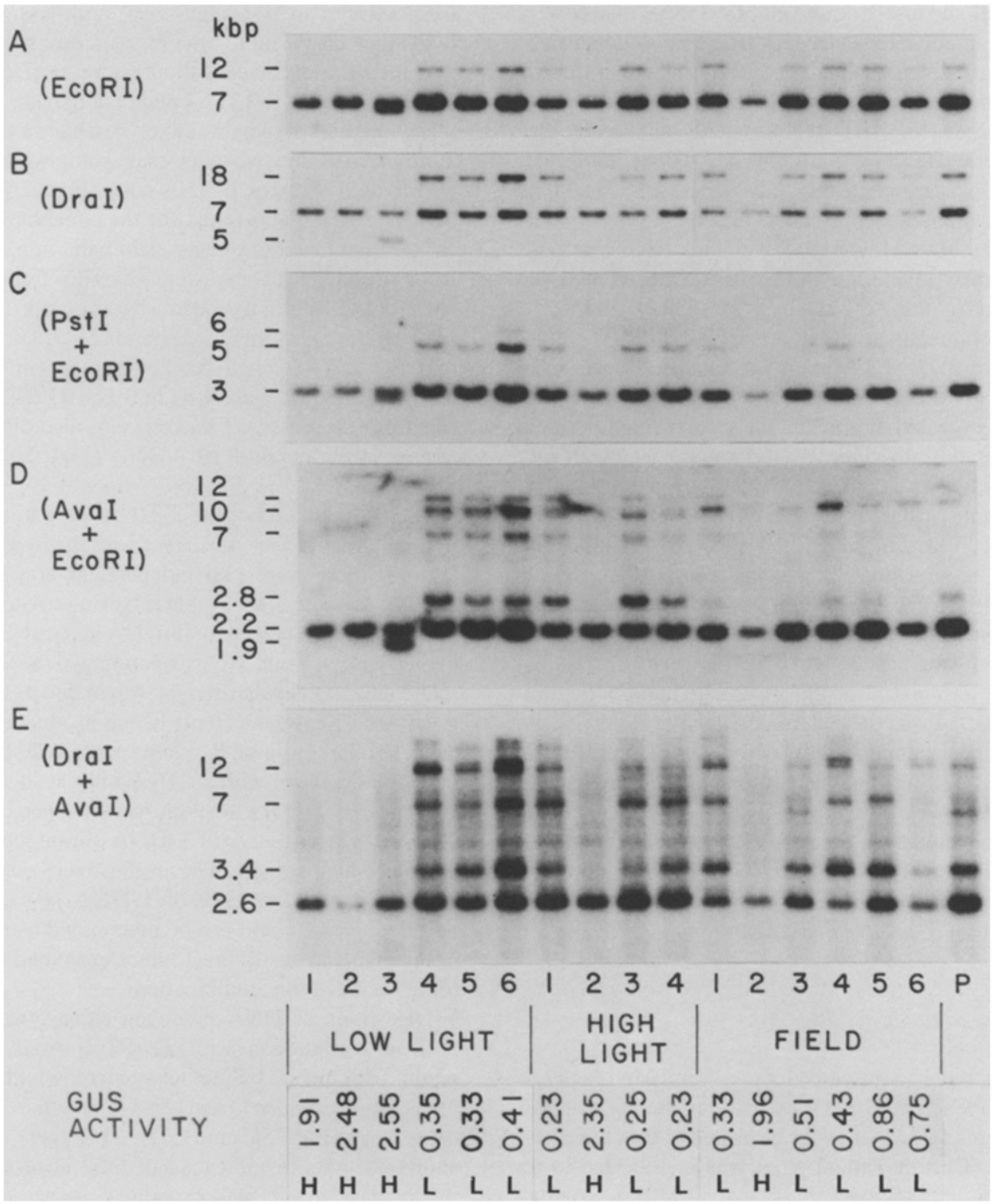

Fig. 5. Autoradiograms of Southern blots of DNA extracted from 16 segregating T3 : L3 R_2 plants grown under different environments and their R_1 parent (P). A to D, Filters hybridised with *uidA* probe and DNA digested with: *Eco*RI (A); *Dra*I (B); *Eco*RI together with *Pst*I (C); and *Eco*RI together with *Ava*I (D). E, DNA digested with *Dra*I together with *Ava*I and filter probed with *kan*. Leaf GUS activity is in nmol cm^{-2} min^{-1} and H or L indicates that plants have high or low GUS expression, respectively. GUS activity of P (from single R_1 determination) was 0.52 nmol cm^{-2} min^{-1}.

methylation at the *Ava*I site in the CaMV-35S promoter, producing only 2.6 kb fragments (Fig. 5E). In comparison, all plants with low GUS expression showed partial methylation at this site, as evidenced by the presence of higher molecular weight fragments (Fig. 5E).

It is possible that only locus 2 T-DNA is methylated and that locus 1 remains unmethylated. Similarly, if T-DNA at locus 2 were not full-length then spurious high molecular weight bands could be produced that could be confused with increased methylation. However, neither of these possibilities seems to be the case as the 7 kb fragments associated with locus 1 appeared in DNA from the low GUS activity plants digested with *Eco*RI and *Dra*I singly or together with *Ava*I (Figs. 5A, 5B, 5D, 5E). This indicates that locus 1 T-DNA has not been fully digested at any internal *Ava*I sites.

Discussion

Variability was found in GUS activity among the R_1 and R_2 generations of 10 tobacco transformants produced by the insertion of T-DNA from pBI121. Replication and synchronisation of R_2 material gave precise measurements of GUS activities in individual transformants. Where T-DNA had been inserted at a single locus, inter-transformant variability was discontinuous, transformants having either high or low GUS activity. High and low groupings were also apparent for individual R_2 plants in one line from T3 which had two independently segregating T-DNA insertions.

In homozygous R_2 material, the high activity group (T5, T13, T14, and T19) had one full-length copy of T-DNA whereas the low activity group (T4, T7, and T18) had double insertions at the same locus, at least one of which was full-length. This negative relationship between number of inserts and GUS activity was also found for the apparently unlinked double T-DNA insertions found in T3. In addition, T6, whose genetic make-up could not be discerned in the present study, also had low GUS activity and multiple copies of T-DNA inserted at several loci. There have been many reports of copy number and gene expression not being linked [4, 15, 16, 19, 27, 29, 30] but our work clearly shows a negative association between number of inserts and expression, something only alluded to previously [16, 27, 29]. These results are similar to those of Napoli *et al.* [25] and van der Krol *et al.* [33] who showed that the introduction into petunia of an extra copy of its own chalcone synthase or dihydroflavonol-4-reductase gene did not result in overexpression of the gene product but in fact could result in a dramatic reduction in gene expression.

For homozygous transformants, there was no indication of position effects within either the high or the low GUS activity group, even though T-DNA had apparently been integrated at different places within the genome. Of necessity, only a comparatively small number of transformants could be examined in detail here. However, as there was great similarity in GUS activity within each group and as the obvious difference between groups was in number of insertions, it is indicated that position effects may not generally be a major cause of inter-transformant variability.

The apparent contradiction that two allelic copies of T-DNA give double expression in the high group (also reported by others [3, 4, 19]) whereas two non-allelic copies, either closely linked or unlinked, effectively reduced expression markedly cannot as yet be explained. Matzke *et al.* [22] showed that the expression of the genes on one T-DNA insert could be completely suppressed by the introduction of another, different T-DNA by retransformation of previously transformed material but our results are from multiple copies of identical T-DNA inserted during a single transformation experiment.

Estimates of copy number in T4 and T7 differed depending on the restriction enzyme used. This can be explained if closely linked inverted repeats were present: *Dra*I-*kan-uidA-Eco*RI: *Eco*RI-*uidA-kan-Dra*I in T4 would give 1 fragment bybridising to *uidA* when cut with *Dra*I and 2 when cut with *Eco*RI; similarly, *Eco*RI-*uidA-kan-Dra*I:*Dra*I-*uidA-kan-Eco*RI in T7 would give 2 fragments when cut with *Dra*I and 1 when

cut with *Eco*RI. This hypothesis is supported by the appearance of a single fragment (approximately 4.5 kb) hybridisable to *uidA* in T4 DNA digested with *Xba*I which cuts uniquely within the T-DNA, between CaMV-35S and *uidA* (Fig. 1A) and 2 fragments in T7 DNA (data not presented).

In general, there was a greater degree of methylation associated with low than with high GUS activity, although all plants with high GUS activity could not be separated from all those with low activity according to the state of methylation of any single site examined in all the cells in the leaf sample. However, all homozygous transformants in the low group showed methylation at the *Ava*I site between *uidA* and its *nos* terminator, which was not methylated in the high group (Fig. 3E). Similarly, all T3 : L3 plants in the low group showed a degree of methylation at the *Ava*I site in the CaMV-35S which was not methylated in the high group (Fig. 5E). Matzke *et al.* [22] also reported that the suppression of genes in the T-DNA, caused by the presence of another insert, was associated with increased methylation within the promoter. In our case, a link between low expression and methylation at any particular site could not be conclusively made. However, in work with the adenovirus major late promoter [34] it was shown that methylation at one nucleotide reduced transcription markedly whereas methylation at another, only 6 base pairs away, had no demonstrable effect. It is possible, therefore, that in our material the key methylation site has yet to be determined. As found by Matzke *et al.* [22], we have been unable to reverse methylation and increase gene expression by treating seedlings with 5-azacytidine (data not presented).

The RNA amounts hybridisable to the *uidA* probe generally correlated with GUS activities for the different transformants. As previously found [15, 20], this indicates that differences in the expression of *uidA* were associated with differences in the level of transcript rather than differences in translation or stability of the reporter gene in a heterologous system.

There were no apparent abnormalities in the transformation protocol used here. Similar percentages of single and multiple insertions at one or more loci were obtained to those reported for other transformation experiments [2, 3, 10].

The experiments reported here used data from R_1 and R_2 generations, hence eliminating the possibility of interference from R_0 epigenetic effects. Such effects have been shown to be a major cause of inter-transformant variability [16] and cannot be excluded as a cause of inter-transformant variability where R_0 material was examined [3, 8, 15, 16, 20, 24, 27, 29, 30, 35]. Some reports showed that subsequent generations had similar levels of variability [4, 27] but others showed that expression in progeny could differ from that in the parents [3] and that non-Mendelian segregation for the expression of the introduced gene can occur [2, 7] especially in plants with high copy number T-DNA [7].

Levels of GUS activity were very similar for comparable plants grown in different environments (Fig. 2) and further duplicated field samples taken 28 days later showed similar results (data not presented). This confirms previous comparisons, using restricted numbers of transformants, of the expression of genes driven by the CaMV-35S in the field, greenhouse and environmental chamber [5, 26].

Not all instances of duplicate T-DNA within a plant result in reduced expression. Jones *et al.* [16] showed that inverted repeats did not preclude high-level expression. Matzke *et al.* [22] noted that the effect only occurred in 50% of plants with two T-DNA copies. In this report one of the T3 : L3 plants (plant 3 from the low-light environment) also had high GUS activity and two T-DNA copies, one of which showed a partial deletion. Further research may indicate whether particular combinations of T-DNA are important in the production of this effect.

Whatever its cause, the fact that double T-DNA insertions can give lower gene expression than single insertions is important for several practical reasons. First, it might explain an apparent loss of expression from generation to generation or an apparent lack of Mendelian inheritance in transformants [2, 7, 10]. Second, it

means that care must be taken in comparing transformants with different promoters, or partial promoter deletions, to ensure that similar material is compared (e.g. single inserts with single inserts). Third, the blue precipitate formed by GUS acting on 5-bromo-4-chloro-3-indolyl β-D-glucuronic acid [13] is often used to detect transformed material. However, it was noted here (data not presented) that leaves from the low GUS expression group (as determined by the MUG assay), whether from old plants or newly germinated seedlings, would rarely if ever show blue coloration with this less sensitive assay (leaves from the high activity group rapidly stained blue throughout). Finally, as different vector systems have been reported to produce differing numbers of inserted genes [17, 27], the results presented here indicate that a system that is capable of constantly introducing single copies might yield more active transformants than one that introduces a large number of inserts.

Acknowledgements

The authors would like to thank Mr. Dion Durnford for his excellent technical assistance and Drs. John Mahon and Lawrence Pelcher for partial sequencing of the T-DNA from pBI121.

References

1. Benfey PN, Chau N-H: Regulated genes in transgenic plants. Science 244: 174–188 (1989).
2. Budar F, Thia-toong L, van Montagu M, Heralsteens J-P: *Agrobacterium*-mediated gene transfer results mainly in transgenic plants transmitting T-DNA as a single mendelian factor. Genetics 114: 303–313 (1986).
3. Czernilofsky AP, Hain R, Baker B, Wirtz U: Studies of the structure and functional organization of foreign DNA integrated into the genome of *Nicotiana tabacum*. DNA 5: 473–482 (1986).
4. Dean C, Jones J, Favreau M, Dunsmuir P, Bedbrook J: Influence of flanking sequences on variability in expression levels of an introduced gene in transgenic tobacco plants. Nucl Acids Res 16: 9267–9283 (1988).
5. De Greef W, Delon R, De Block M, Leemans J, Botterman J: Evaluation of herbicide resistance in transgenic crops under field conditions. Bio/technology 7: 61–64 (1989).
6. Dellaporta SL, Wood J, Hicks JB: Maize DNA miniprep. In: Malmberg R, Messing J, Sussex I (eds) Molecular Biology of Plants: a Laboratory Course Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1985).
7. Deroles SC, Gardner RC: Analysis of the T-DNA structure in a large number of transgenic petunias generated by *Agrobacterium*-mediated transformation. Plant Mol Biol 11: 365–377 (1988).
8. Eckes P, Rosahl S, Schell J, Willmitzer L: Isolation and characterization of a light inducible, organ-specific gene from potato and analysis of its expression after tagging and transfer into tobacco and potato shoots. Mol Gen Genet 205: 14–22 (1986).
9. Gamborg OL, Miller RA, Ojima K: Nutrient requirements of suspension cultures of soybean root cells. Exp Cell Res 50: 151–158 (1968).
10. Heberle-Bors E, Charvat B, Thompson D, Schernthaner JP, Barta A, Matzke AJM, Matzke MA: Genetic analysis of T-DNA insertions into the tobacco genome. Plant Cell Rep 7: 571–574 (1988).
11. Hobbs SLA, Pelcher LE, DeLong CMO, Anderson M, Mahon JD: Genetic variability in the amount of ribulose-1, 5-bisphosphate carboxylase/oxygenase and its small subunit mRNA in pea. Planta 180: 510–516 (1990).
12. Horsch RB, Fry JE, Hoffman NL, Eichholz D, Rogers SG, Fraley RT: A simple and general method for transferring genes into plants. Science 27: 1229–1231 (1985).
13. Jefferson RA, Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol Biol Rep 5: 387–405 (1987).
14. Jefferson RA, Kavanagh TA, Bevan MW: GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6: 3901–3907 (1987).
15. Jones JDG, Dunsmuir P, Bedbrook J: High level expression of introduced chimaeric genes in regenerated transformed plants. EMBO J 4: 2411–2418 (1985).
16. Jones JDG, Gilbert DE, Grady KL, Jorgensen RA: T-DNA structure and gene expression in petunia plants transformed by *Agrobacterium tumefaciens* C58 derivatives. Mol Gen Genet 207: 478–485 (1987).
17. Jorgensen R, Snyder C, Jones JDG: T-DNA is organized predominantly in inverted repeat structures in plants transformed with *Agrobacterium tumefaciens* C58 derivatives. Mol Gen Genet 207: 471–477 (1987).
18. Koncz C, Schell J: The promoter of T_L-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. Mol Gen Genet 204: 383–396 (1986).
19. Last DI, Gray JC, Synthesis and accumulation of pea plastocyanin in transgenic tobacco plants. Plant Mol Biol 14: 229–238 (1990).
20. Lawton MA, Tierney MA, Nakamura I, Anderson E, Komeda Y, Dube P, Hoffman N, Fraley RT, Beachy RN: Expression of a soybean β-conglycinin gene under the control of the cauliflower mosaic virus 35S and 19S

promoters in transformed petunia tissues. Plant Mol Biol 9: 315–324 (1987).
21. Le Gal MF, Hobbs SLA, DeLong CMO: Gene expression during the infection process in nodulating and non-nodulating pea genotypes. Can J Bot 67: 2535–2538 (1989).
22. Matzke MA, Primig M, Trnovshy J, Matzke AJM: Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants. EMBO J 8: 643–649 (1989).
23. Murashige T, Skoog F: A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 15: 473–497 (1962).
24. Nagy F, Odell JT, Morelli G, Chua N-H: Properties of expression of the 35S promoter from CaMV in transgenic tobacco plants. In: Zaitling M, Day P, Hollaender A (eds) Biotechnology in Plant Science. Academic Press Inc, London (1985).
25. Napoli C, Lemieux C, Jorgensen R: Introduction of a chimeric chalcone synthase gene into petunia results in reversible cosuppression of homologous genes in *trans*. Plant Cell 2: 279–289 (1990).
26. Nelson RS, McCormick SM, Delannay X, Dube P, Layton J, Anderson EJ, Kaniewska M, Proksch RK, Horsch RB, Rogers SG, Fraley RT, Beachy RN: Virus tolerance, plant growth, and field performance of transgenic tomato plants expressing coat protein from tobacco mosaic virus. Bio/technology 6: 403–409 (1988).
27. Odell JT, Nagy F, Chua N-H: Variability in 35S promoter expression between independent transformants. In: Key J, McIntosh L (eds) Plant Gene Systems and their Biology, vol 62. Alan R. Liss, New York (1987).
28. Potrykus I, Paskowski J, Saul M, Petruska J, Shillito RD: Molecular and general genetics of a hybrid foreign gene introduced into tobacco by direct gene transfer. Mol Gen Genet 199: 169–177 (1985).
29. Rogers SG, O'Connell K, Horsch RB, Fraley RT: Investigation of factors involved in foreign protein expression in transformed plants. In: Zaitlin M, Day P, Hollaender A (eds) Biotechnology in Plant Science. Academic Press Inc, London (1985).
30. Shirsat AH, Wilford N, Croy RRD: Gene copy number and levels of expression in transgenic plants of a seed specific gene. Plant Sci 61: 75–80 (1989).
31. Snedecor GW, Cochran WG: Statistical Methods. The Iowa State University Press, Ames, Iowa (1980).
32. Spielmann A, Simpson RB: T-DNA structure in transgenic tobacco plants with multiple independent integration sites. Mol Gen Genet 205: 34–41 (1986).
33. van der Krol AR, Mur LA, Beld M, Mol JNM, Stuitje AR: Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. Plant Cell 2: 291–299 (1990).
34. Watt F, Molloy PL: Cytosine methylation prevents binding to DNA of an HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. Genes Dev 2: 1136–1143 (1988).
35. Williamson JD, Hirsch-Wyncott ME, Larkins BA, Gelvin SB: Differential accumulation of a transcript driven by the CaMV 35S promoter in transgenic tobacco. Plant Physiol 90: 1570–1576 (1989).