Mol Gen Genet (1987) 207:478–485

MGG

T-DNA structure and gene expression in petunia plants transformed by *Agrobacterium tumefaciens* C58 derivatives

Jonathan D.G. Jones, David E. Gilbert, Karen L. Grady*, and Richard A. Jorgensen
Advanced Genetic Sciences, Inc., 6701 San Pablo Ave., Oakland, CA 94608, USA

Summary. We have previously described substantial variation in the level of expression of two linked genes which were introduced into transgenic petunia plants using *Agrobacterium tumefaciens*. These genes were (i) nopaline synthase (*nos*) and (ii) a chimeric chlorophyll a/b binding protein/octopine synthase (*cab*/*ocs*) gene. In this report we analyze the relationship between the level of expression of the introduced genes and T-DNA structure and copy number in 40 transgenic petunia plants derived from 26 transformed calli. Multiple shoots were regenerated from 8 of these calli and in only 6 cases were multiple regenerated shoots from each callus genotypically identical to each other. Many genotypes showed no *nos* gene expression (22/28). Most of the plants (16/22) which lacked *nos* gene expression did contain *nos*-encoding DNA with the expected restriction enzyme map. Similarly, amongst the genotypes showing no *cab*/*ocs* gene expression, the majority (11/28) did not show any alterations in restriction fragments corresponding to the expected *cab*/*ocs* coding sequences (10/11). Approximately half of the plants carried multiple copies of T-DNA in inverted repeats about the left or right T-DNA boundaries. No positive correlation was observed between the copy number of the introduced DNA and the level of expression of the introduced genes. However, plants with high copy number complex insertions composed of multiple inverted repeats in linear arrays usually showed low levels of expression of the introduced genes.

Key words: *Agrobacterium tumefaciens* – Transgenic plants – T-DNA structure – Between-transformant variability – Chimeric genes

Introduction

In a previous study (Jones et al. 1985) we investigated the levels of expression of two linked introduced genes in transgenic petunia plants derived from cocultivation of petunia protoplasts with various engineered *Agrobacterium tumefaciens* strains. These strains were made using pGV3850, a non-oncogenic Ti plasmid derivative of C58 (Zambryski et al. 1983). The introduced genes consisted of a wild-type nopaline synthase (*nos*) gene and a fusion between a petunia chlorophyll a/b binding protein (*cab*) gene and the octopine synthase (*ocs*) gene. Substantial between-transformant variation was observed in the level of expression of both genes. Some plants showed uncoordinated variability of the chimeric *cab*/*ocs* and *nos* genes, i.e., they either had an Ocs^+ Nos^- or an Ocs^- Nos^+ phenotype. Furthermore, substantial variation was sometimes observed between shoots regenerated from the same transformed callus.

It is possible that some of this variability in T-DNA gene expression is due to the different positions at which the T-DNA has integrated into the genome of different transformants. However, a more easily tested hypothesis is that the variation can be explained by differences in the copy number of the inserted DNA in each plant, by DNA rearrangements, by deletions or by incomplete transfers of the T-DNA which preclude expression of affected genes. Also, in the accompanying paper (Jorgensen et al. 1987) we show that multiple copies of T-DNA frequently integrate in the form of inverted repeats and so we were interested in whether this structure was in any way correlated with levels of gene expression.

We describe here experiments which characterize the structure of the T-DNA in 40 transgenic petunia plants derived from 26 kanamycin-resistant calli. DNA from each of these plants was subjected to Southern analysis using probes homologous to one of four different locations in the T-DNA. The structure and organization of the T-DNA in each plant was defined with respect to these probes in terms of copy number and rearrangements. In particular, characterization of the ends of the T-DNAs was conducted in order to determine whether or not multiple T-DNA elements were joined as inverted or tandem repeats. The data obtained permit an evaluation of the extent to which the copy number and structure of the T-DNA rather than its location in the genome can explain the variation in the levels of expression of the introduced genes.

Materials and methods

Genetic materials. The transformed petunia plants analyzed here have been described (Jones et al. 1985) except for those individuals transformed with C58C1/pGV3850 kanR::502. This cointegrate is formed with a plasmid which differs from 503 only in the fact that the orientation of the *cab*/*ocs* fusion in the pMUC9 plasmid is reversed. A map of the T-DNA region of the cointegrate *Agrobacterium* strain carrying plasmid pGV3850 kanR::511 is shown in Fig. 1F.

* *Present address:* Department of Biology, College of Charleston, Charleston, SC 29424, USA

Offprint requests to: J. Jones

Blot hybridization analysis. DNA was prepared from petunia plants using a modification of the CTAB procedure of Murray and Thompson (1980). Southern blot analysis was performed essentially according to Southern (1975) and Maniatis et al. (1982). Five micrograms of digested plant DNA was loaded in each lane. Probes were made using SP6 RNA polymerase and [^{32}P]UTP (600 Ci/mmol, Amersham) as described by the manufacturer (ProMega Biotech). Hybridizations were carried out at 42° C in 50% formamide, 10% dextran sulfate and 5 × SSC. Prehybridizations were carried out for 24 h and hybridizations for 36 h. Filters were washed in 2 × SSC, 0.1% SDS at 42° C and then at 65° C in 0.1 × SSC, 0.5% SDS.

Construction of probe plasmids. pKL133 contains a 3.0 kb *Hin*dIII-*Eco*RI fragment from the T-DNA part of the *Hin*dIII–10 fragment of the C58 Ti plasmid (Joos et al. 1983) cloned into pSP64. It was linearized with *Eco*RI prior to SP6 RNA polymerase transcription reactions. pJJ1124 consists of a 1.0 kb *Bgl*II-*Sma*I fragment carrying the *nptII* gene of Tn*5* cloned into pSP64 cut with *Hin*cII and *Bam*HI and was linearized with *Eco*RI prior to use. pJJ1019 has been described (Jones et al. 1985). pJJ1104 consists of a 1.1 kb *Cla*I-*Pvu*II fragment homologous to nopaline synthase coding sequences (from +455 to +1550 in Depicker et al. 1982) cloned into pSP64 cut with *Acc*I and *Sma*I. It was linearized with *Eco*RI prior to use. Restriction enzyme digestions, fragment purifications, ligation, transformations of *Escherichia coli* and plasmid purifications were carried out essentially as described by Maniatis et al. (1982).

Results

Transgenic plants for which the T-DNA structure was characterized

We describe here the analysis of T-DNA structures in 40 transgenic petunia plants containing *cab/ocs* fusions. These fusions were introduced as cointegrates with the T-DNA of the pGV3850 derivative pGV3850 kanR. The *cab/ocs* fusion constructions (503, 511, 521, 571) have been described (Jones et al. 1985). The plasmids 511 and 521 contained *cab/ocs* fusions with a junction within the leader sequence (transcriptional fusions) while the plasmids 503 and 571 contained *cab/ocs* fusions with a junction in frame and within the coding sequence (translational fusion). We also analyzed DNA from three plants transformed with construction 502 in which the *cab/ocs* gene is in the opposite orientation to 503 (see Materials and methods). After cocultivation, individual kanamycin-resistant calli were selected and regenerated into whole plants. Each transformant was numbered to indicate the construction from which it was derived (e.g. 571.7) and where multiple shoots were regenerated from the same callus each shoot was designated by a letter (e.g., 571.7A, 571.7B, 571.7C). Of the 52 plants described in Jones et al. (1985), 37 are analyzed here (see Table 2). The original 52 plants were derived from 27 independent calli, shoots from 23 of which are investigated here.

Strategy for the analysis of T-DNA structures

Southern hybridizations were conducted in order to analyze in detail the structures of the integrated T-DNAs. Four different single-stranded RNA probes were hybridized to DNA from the transformed plants (Fig. 1F, and Materials and methods). These were: (i) pJJ1104 homologous to T-DNA near the right border; (ii) pJJ1019 homologous to the octopine synthase gene (used as a probe in the Northern hybridizations of Jones et al. 1985); (iii) pJJ1124 homologous to the neomycin phosphotransferase II (*npt*) gene of Tn*5*; and (iv) a probe pKL133 homologous to T-DNA near the left border. Probe pJJ1019 carries sequences from the 3′ untranslated region of the *ocs* gene and consequently displays some homology to the chimeric p_{nos}/*npt*/*ocs* 3′ gene which was used as a selectable marker.

Table 1. DNA digests to which each probe was hybridized

Probe	pKL133	pJJ1124	pJJ1019	pJJ1104
Digests	*Hin*dIII *Eco*RI *Bgl*II + *Xba*I *Bgl*II	*Hin*dIII *Nco*I + *Eco*RI	*Nco*I	*Hin*dIII *Hin*dIII + *Bam*HI *Hin*dIII + *Bcl*I

Table 1 tabulates the DNA digests to which each probe was hybridized. In all cases hybridizations were carried out with positive and negative controls. The negative control was DNA from untransformed petunia and the positive controls were one-, four- and ten-copy reconstructions using purified plasmid DNA and DNA from the *Agrobacterium* strains (C58/pGV3850 kanR::502, 503, 511, 521 or 571) which had been used to transform the petunia cells.

The left and right boundary region probes and associated digests provide criteria for recognizing the presence of inverted repeats (IRs) at the ends of the T-DNA. In principle, an IR about a T-DNA end will give border fragments of a size twice the distance from the border to the nearest recognition site for a particular enzyme. Also, bands corresponding to such fragments should hybridize with twice, or an integral multiple of twice, the signal of a band from a non-inverted repeat border fragment. In the event of an IR at the right border, *Hin*dIII or *Hin*dIII + *Bam*HI digests of the T-DNA yield diagnostic fragments when a right border probe is used (Fig. 1F and Jorgensen et al. 1987). A *Hin*dIII site lies about 2.25 kb from the right border and an approximately 4.5 kb fragment is frequently observed when a right border probe is used on blots of *Hin*dIII digests (Fig. 1C). A *Bam*HI site lies midway between the *Hin*dIII site and the right border, and when IRs are present *Hin*dIII + *Bam*HI digests give 1.1 kb and 2.3 kb fragments homologous to pJJ1104 (Jorgensen et al. 1987). In the Ti plasmid, a *Bcl*I site lies 39 bp to the left of the 25 bp repeat which marks the right end of the T-DNA. Consequently, the right border in the T-DNA of all plants was also analyzed by digestion with *Hin*dIII + *Bcl*I in order to investigate whether the transfer of this *Bcl*I site had occurred. It is conceivable that the occurrence of junctions to the left of this site could render the *nos* promoter inactive since the *nos* transcriptional initiation site is only 306 bp from the right repeat (Depicker et al. 1982).

The left boundary probe pKL133 is defined at its right end by a *Hin*dIII site which is only 29 bp from the next *Eco*RI site. These two sites lie about 3.0 kb to the right of the left repeat. In a left border IR, *Hin*dIII and *Eco*RI digests should lead to pKL133-homologous fragments which are almost identical in size (though the actual size is more variable since the position of the left end of the

T-DNA is more variable). The only exception to this is when the *Eco*RI site 52 bp to the right of the left repeat is successfully transferred to the plant nucleus. This is a rare event (Jorgensen et al. 1987). When the *Eco*RI and *Hin*dIII digests taken together suggest an IR, the sizes of a *Bgl*II fragment and *Bgl*II + *Xba*I fragments also homologous to this probe can be predicted and tested to confirm an IR identification.

Internal probes were hybridized to appropriate digests to test for both the presence and integrity of the internal fragments. At issue here is whether the *cab/ocs* null phenotype which is seen in some of the transgenotes results in some or all cases from deletion or rearrangement of the *cab/ocs* coding sequences. Additionally, data from these hybridizations permit an assessment of the T-DNA copy number based on band intensity. This can be compared with a copy number estimate based on the number of left and right border fragments.

Blot hybridization analysis of the structures of different T-DNAs

Figure 1 presents five examples of autoradiographs of representative Southern hybridizations. A map of the cointegrate formed with construction 511 is also included as an example of a restriction enzyme map of the cointegrates (Fig. 1F).

Figure 1A and B shows hybridizations using the left border probe pKL133 to blots of *Eco*RI + *Nco*I and *Hin*dIII digests, respectively, of DNA from all of the 502 and 503 plants and one 511 plant. Three conclusions can be made on the basis of these data and on the basis of data derived from hybridizations with this probe to the other DNA samples.

Firstly, some transformants, e.g., 503.1B (lane 6), 503.5A, B, C, and D (lanes 9, 10, 11, 12) give rise to simple patterns consistent with only one left border fragment while others, e.g., 502.3A (lane 3), 502.6A (lane 4), 503.1A (lane 5), and 503.9A (lane 15), show complex patterns with multiple bands. However, plants arising from different transformed calli always show a unique fragment or set of fragments.

Secondly, shoots regenerated from the same callus usually show the same border fragments, e.g., 503.5A, B, C, D (lanes 9, 10, 11, 12); 503.9A , 9C (lanes 15, 16); and 503.10A, 10B (lanes 17, 18). However, in one example two shoots from the same callus show different border fragments (503.1A, 1B, lanes 5, 6). The same is true (data not shown) for 511.2A and 511.2B. On this basis we concluded that the 40 different plants which arose from 26 calli were probably derived from 28 independent transformation events.

Thirdly, fragments are observed which are consistent with the existence of inverted repeats using the criteria defined in the previous section. Bands in DNA from the same individual which are the same size in *Eco*RI and *Hin*dIII digests are indicated by an arrow. On the basis of bands observed in blots of *Bgl*II and *Bgl*II + *Xba*I digests (not shown), it was concluded that the arrowed bands in Fig. 1A and B were due to inverted repeats about the left border.

A summary of the interpretations of all of the patterns observed in all the blots is given in Table 2. In the "Left border" columns of the table we present our conclusions concerning the number of T-DNA copies which either are (IR) or are not (NIR) involved in inverted repeats at the left border.

Figure 1C shows *Hin*dIII digests of the same DNA samples hybridized with a right border probe. These digests, together with results from *Hin*dIII + *Bam*HI and *Hin*dIII + *Bcl*I double digests (data not shown) permit an analysis of the right border of the T-DNA. Again, Table 2 shows an interpretation of the structures of the right borders of the T-DNA of the various independent transformants. It should be noted that DNA from 503.1B, 511.6A, and 571.10A, B and C, plants which show no nopaline synthase expression (see later), exhibits no homology to a right border probe. In the "Right border" columns in Table 2 designated "Bcl", individual plants are assessed for the number of right border copies they contain of a *Hin*dIII + *Bcl*I fragment of the expected 2.25 kb size (column "exp") or of some other size (designated as "not exp").

Figure 1D shows a *Hin*dIII digest hybridized with pJJ1124. There is an internal *Hin*dIII fragment in the DNA of all transformants whose intensity reflects the overall copy number of the introduced DNA. In one lane (511.2A, lane 19) a band of 2.8 kb, smaller than the 4.8 kb size predicted on the basis of the map (Fig. 1F), was observed. Consequently, in Table 2 this was entered in the "not exp" column.

Figure 1E shows similar data for seven lanes of *Nco*I-digested DNA hybridized with the pJJ1019 probe. These data were generated in order to confirm the presence and integrity of the *cab/ocs* gene. Lane 1 is an untransformed DNA negative control lane. An arrow draws attention to aberrantly migrating bands in 511.6A (lane 3), 511.8A (lane 4) and 521.1B (lane 5) compared with the expected sizes shown in lanes carrying DNA from 511.2B (lane 2), 521.2A (lane 6) and 521.4A (lane 7). Interestingly, 521.2A is the plant shown by Jones et al. (1985) to be Nos^+ and Ocs^-. In this case the Ocs^- phenotype does not appear to be due to rearrangement of the coding sequence. These data and data for the other transformants (not shown) were incorporated into Table 2. Plants were evaluated with respect to T-DNA copy number (based on relative band intensity) and whether or not correctly migrating bands were observed. The question of the origin of the aberrantly migrating bands was not pursued.

Interpretation of T-DNA structure and copy number

Our interpretations of the copy number, organization and integrity of fragments hybridizing to each probe in DNA from each transgenote are summarized in Table 2. Due to the large number of DNA samples which we analyzed a less comprehensive analysis is reported here than that conducted on 11 tomato transformants by Jorgensen et al. (1987) in that a smaller number of combinations of probes and digests were used on the left and right borders. Also, some of the hybridization patterns observed are more complex and more difficult to interpret. However, the model for each T-DNA is consistent with the results of ten different hybridizations of specific probes to specific digests and we are confident in each case that we have proposed the most likely structure. The complexity of certain integration events means that some bands can be explained in more than one way and so we may have under- or overestimated the numbers of inverted or direct repeats in some plants. Seven classes of sequence organization pattern were ob-

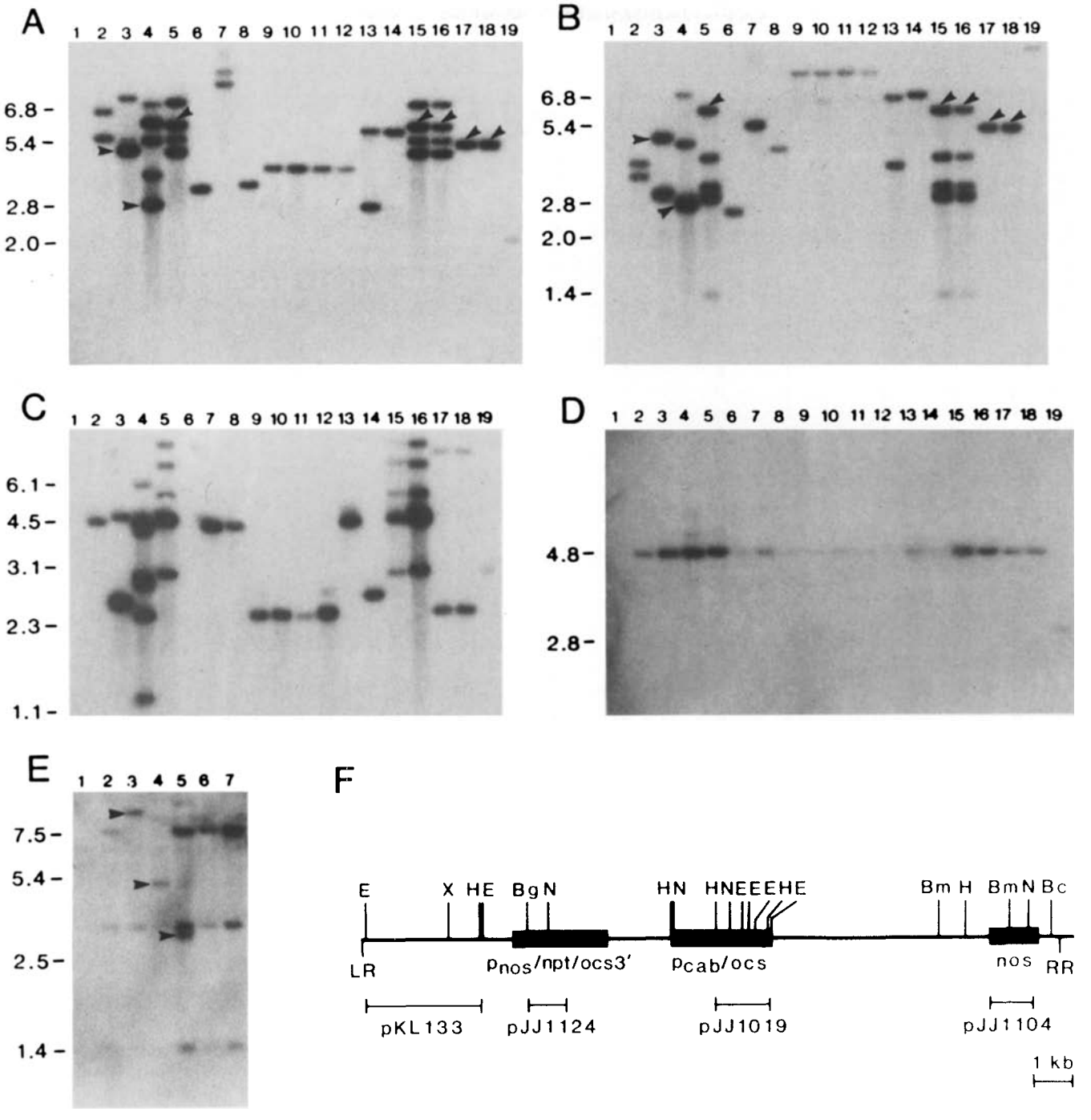

Fig. 1A–F. Autoradiographs and maps of T-DNA structures. **A**, **B**, **C** and **D** show autoradiographs resulting from Southern hybridizations using four different combinations of probe and digest with the same 19 DNA samples. At the side of each panel the sizes of comigrating size marker fragments are indicated. In each panel the lanes are: (1) untransformed petunia DNA, (2) 502.2B, (3) 502.3A, (4) 502.6A, (5) 503.1A, (6) 503.1B, (7) 503.2B, (8) 503.3A, (9) 503.5A, (10) 503.5B, (11) 503.5C, (12) 503.5D, (13) 503.6A, (14) 503.7A, (15) 503.9A, (16) 503.9C, (17) 503.10A, (18) 503.10B and (19) 511.2A **A** *Hin*dIII digests probed with pKL133. **B** *Eco*RI+*Nco*I digests probed with pKL133. The arrows in **A** and **B** indicate bands in the DNA of particular transformants which hybridize to pKL133 and which are the same size in both *Hin*dIII and *Eco*RI digests (**F** indicates the location of homology to pKL133). **C** *Hin*dIII digests hybridized with pJJ1104. The 4.5 kb size class is the one which corresponds to that observed in the most common class of inverted repeat. **D** *Hin*dIII digests hybridized with pJJ1124. An aberrantly migrating band is observed in 511.2A (lane 19). **E** *Nco*I digests hybridized with pJJ1019. The lanes are: (1) DNA from untransformed petunia plants, (2) 511.2B, (3) 511.6A, (4) 511.8A, (5) 521.1B, (6) 521.2A and (7) 521.6A. Lanes 2, 5, 6 and 7 show the expected bands of 7.5 kb, 3.0 kb (due to the *ocs* 3′ end on the *nos/nptII* gene) and 1.4 kb. Lanes 3, 4 and 5 show fragments which migrate with other than the expected size (*arrowed*). **F** shows a restriction map of a typical cointegrate, GV3850 kanR::511. The symbols are: LR, left repeat; RR, right repeat; E, *Eco*RI; X, *Xba*I; H, *Hin*dIII; Bg, *Bgl*II; N, *Nco*I; Bm, *Bam*HI; Bc, *Bcl*I. Regions of DNA homologous to each probe are indicated under the figure except for the homology between the *ocs* probe pJJ1019 and the *ocs* 3′ end of the *nos/nptII* gene

Table 2. Summary of interpretations of T-DNA structures in different transgenotes

Transgenote	Left border		*nos/npt*		*cab/ocs*		Right border				Expression	
	IR	NIR	exp	not exp	exp	not exp	IR	NIR	Bcl exp	Bcl not exp	*ocs*	*nos*
502 2B	0	2	2	0	2	0	2	0	2	–	0	–
502 3A	4	1	5	0	5	0	2	4	1	4	4	–
502 6A	6	2	10	0	10	0	6	3	8	1	0	–
503 1A	6	1	8	0	4	3	6	1	5	3	0.5	–
503 1B	0	1	1	0	1	0	Absent		Absent		100	–
503 2B	0	2	2	0	2	0	2	0	2	0	0.5	–
503 3A	0	1	1	0	1	0	0	1	1	0	85	+
503 5A	0	1	1	0	1	0	0	1	0	1	50	–
503 5B	0	1	1	0	1	0	0	1	0	1	6	–
503 5C	0	1	1	0	1	0	0	1	0	1	20	–
503 5D	0	1	1	0	1	0	0	1	0	1	12	–
503 6A	0	2	2	0	2	0	2	0	2	0	0	–
503 7A	0	1	1	0	1	0	0	1	1	0	10	+
503 9A	4	3	7	0	4	3	2	5	5	2	0	–
503 9B	4	3	7	0	4	3	2	5	5	2	0	–
503 9C	4	3	7	0	4	3	2	5	5	2	0	–
503 10A	2	0	2	0	2	0	0	2	2	0	6	+
503 10B	2	0	2	0	2	0	0	2	2	0	95	+
511 2A	0	1	0	1	1	0	0	1	1	0	0.5	+
511 2B	0	2	1	0	1	0	0	1	0	1	0	–
511 6A	0	2	1	0	0	1	Absent		Absent		2.0	–
511 8A	0	2	1	0	0	1	0	1	0	1	0	–
521 1B	2	2	4	0	2	2	0	4	2	2	0	–
521 2A	0	2	2	0	2	0	2	2	4	0	0	+
521 4A	4	1	5	0	5	0	4	1	4	0	0	–
521 6A	0	5	5	0	5	0	2	3	5	0	0	–
521 7B	2	2	4	0	4	0	4	0	4	0	6	–
571 3A	0	1	Absent		0	1	0	1	1	0	0.5	–
571 4A	0	2	2	0	2	0	0	2	1	1	35	–
571 5B	0	2	2	0	2	0	2	0	2	0	0	–
571 7A	0	2	2	0	2	0	2	0	2	0	35	–
571 7B	0	2	2	0	2	0	2	0	2	0	75	–
571 7C	0	2	2	0	2	0	2	0	2	0	20	–
571 8A	0	2	2	0	0	2	0	1	2	0	12	–
571 9A	0	1	1	0	1	0	0	1	0	1	12	+
571 9B	0	1	1	0	1	0	0	1	0	1	20	+
571 9C	0	1	1	0	1	0	0	1	0	1	17	+
571 10A	0	2	0	2	0	2	Absent		Absent		17	–
571 10B	0	2	0	2	0	2	Absent		Absent		40	–
571 10C	0	2	0	2	0	2	Absent		Absent		15	–

These interpretations are based on the different DNA digest and probe combinations used in the Southern hybridizations and on the various assays performed to evaluate expression levels of the introduced genes. Analysis of the left and right borders as described in the text and presented in Fig. 1 permits an evaluation of the number of DNA copies involved in inverted repeat (IR) or non-inverted repeat (NIR) junctions. For the internal fragment hybridizations, samples were evaluated by band intensity for the copy number of fragments migrating at both the correct (exp) and incorrect (not exp) sizes. In the right border column, the number of *Hin*dIII + *Bcl*I fragments migrating with the expected size (based on the assumption of complete transfer leftward from the right repeat) was scored in column "Bcl exp" and those migrating at a different size in column "Bcl not exp" (see text). Absence of any detectable homology is scored as "Absent". Levels of gene expression are presented in the "*ocs*" column as a percentage of maximal gene expression observed in 503.1 B while *nos* expression is presented as detectable (+) or not detectable (–)

served and some transformants fell into more than one class.

Seven transformants (25%) contained single copy insertions (503.1B, 503.3A, 503.5A, 503.7A, 511.2A, 571.3A, and 571.9A). Two of these showed less than a complete DNA copy: DNA from 503.1B showed no homology to the *nos* gene (right border) probe and DNA from 571.3A had lost internal homology to the *nptII* gene probe suggesting an internal deletion.

Six transformants (21%) had T-DNA copies arranged in "simple" inverted repeats, five about the right border (502.2B, 503.2B, 503.6A, 571.5B, 571.7A) and one about the left border (503.10A).

Four transformants (14%) exhibited two inserts which were apparently unlinked (511.2B, 571.4A, 571.8A, 571.10A) although the T-DNAs in 571.10A could be organized as an inverted repeat about a right border which was deleted to the left of homology to pJJ1104.

The T-DNA in 511.2B and 571.8A is also representative of a fourth class of insert in which the copy number esti-

mates for the right and left border are apparently different. Other examples of this are 511.6A, 511.8A and 521.2A. This could be due to an inverted repeat which is more difficult to discern, to incomplete insertion or partial deletion of some T-DNA copies or to the undetected presence of direct repeats. It is possible that the copy number estimates based on band intensity are accurate to only plus or minus one copy in the high copy number insertions (see 503.1A in Table 2). For this reason we may be underestimating the frequency of events which give rise to fragments with different copy numbers of left and right border fragments.

Five (17%) of the transformants have four or five T-DNA copies and reasonable interpretations can be put on their structures based on counting the number of borders which are in inverted or non-inverted conformations (502.3A, 521.1B, 521.4A, 521.6A, 521.7B). If a single copy of T-DNA is represented as AZ (for a fragment extending from the left to the right border) then 521.4A and 521.7B can be represented respectively as AZ.ZA.AZ.ZA.AZ and AZ.ZA.AZ.ZA. 502.3A can be represented as ZA.AZ.ZA with two unlinked AZs. Similarly, 521.1B could be of the form ZA.AZ with two unlinked AZs and 521.6A could be of the form AZ.ZA with three unlinked AZs.

Three transformants (502.6A, 503.1A and 503.9A) are so complex that a variety of possible interpretations can be placed on their structures. However we presume that the T-DNAs in these plants are of the form described above involving arrays of various numbers of inverted repeats and also some unlinked single copy insertions.

These categories include all 28 independent transformants. Eight of these transformants were also examples of a seventh class of transformation events which led to T-DNA structures with restriction fragments having unexpected mobilities when hybridized with the internal *npt* or *ocs* probes (503.1A, 503.9A, 511.6A, 511.8A, 521.1B, 571.3A, 571.8A and 571.10A). One of these (511.6A) could simply be due to the same deletion or incomplete transfer event which is revealed by the absence of hybridization with the *nos* probe.

We sought to identify T-DNAs in the form of a direct repeat about a T-DNA end using the criterion of a *Hin*dIII fragment of between 4.5 kb and 5.25 kb (depending on the position of the left border) which hybridizes to both left and right border probes. Our data do not rule out the possibility of direct repeats in the complex integrations of 502.6A, 503.1A, and 503.9A or in the simpler integration of 521.1B. However we are confident in concluding that direct repeats are created less frequently than inverted repeats since DNA from all other transgenotes failed to meet this criterion. This finding has also been reported in the accompanying paper of Jorgensen et al. (1987) after an analysis of 11 tomato transformants.

Expression of the nopaline synthase gene

Data concerning the expression of the nopaline synthase gene in some transformants have already been presented (Jones et al. 1985). Using the same assay nopaline accumulation was investigated in all remaining transformants (data not shown). However, since this assay for gene expression is not fully comparable with the RNA hybridisation assay used in assessing *cab/ocs* gene expression, the same RNA samples previously (Jones et al. 1985) assayed for *cab/ocs* mRNA were also assayed for *nos* mRNA using Northern blots. These data were in complete agreement with the data previously obtained using nopaline accumulation assays. Hybridizations to the same filters with an *ocs* probe both confirmed previous assessments of relative levels of *ocs* mRNA and permitted assessment of the levels of the previously undescribed transgenotes (502 series). Data previously obtained (Jones et al. 1985) and data subsequently obtained by the same techniques reported in Jones et al. (1985) are summarized in Table 2 in the column labelled "Expression". Most of the values in the *ocs* column are taken from Table 1 of Jones et al. (1985). *Nos* gene expression was scored simply as (+) and (−) because the levels of expression were so low even in Nos^+ plants that no meaningful quantitative comparison could be made.

Discussion

Structures of the integrated T-DNAs

We describe here the results of a detailed analysis of the T-DNA in 40 transgenic petunia plants that were derived from 26 calli and show that they are probably the result of 28 independent transformation events. As in the case of the transgenic tomato plants described in the accompanying paper (Jorgensen et al. 1987), more than half (15/28) of these petunia transgenotes exhibit inverted repeats of the T-DNA about either the left border (2 examples), right border (7 examples) or both borders (6 examples). A detailed discussion of the origin and significance of inverted repeat structures can be found in Jorgensen et al. (1987). Only one-quarter (7/28) of the transformation events generated single copy T-DNA inserts. This could have been due to the particular conditions of the cocultivation which involved a high ratio of bacteria to plant cells (50:1)

The data presented in this paper indicate that shoots from the same callus are usually somatic clones. However, there are two exceptions to this rule (503.1A, 1B and 511.2A, 2B). They may have arisen due to clumping of independent transformed calli after cocultivation followed by regeneration of shoots from different sections of the chimeric callus. An alternative hypothesis for the difference between 503.1A and 503.1B would be that the complex structure in 503.1A experienced a large deletion to give the simple structure in 503.1B. It is also conceivable that when a cell is transformed during the period of the cell-cycle between G2 and mitosis, independent T-DNAs transform chromosomes which segregate from each other during the subsequent cell division and thus give rise to a chimeric callus.

T-DNA structure and expression of introduced genes

The data in Table 2 permit several conclusions to be drawn concerning the expression of introduced genes in relation to the T-DNA structure in each transgenic plant. Firstly, there is no simple correspondence between copy number of the inserted DNA and the level of expression of the introduced genes. This is clearly indicated, for example, by comparing 503.1B with 503.1A or 503.9A. 503.1B contains one T-DNA copy but shows high levels of *cab/ocs* expression while 503.1A and 503.9A, which contain multiple copies, half of which show no apparent rearrangement relative to wild type, show very low or undetectable levels of expression of this gene. This is consistent with the data of Jorgensen et al. (1987) who report that a transformant

with five intact copies of the *nos* gene in an array of direct and inverted repeats of the T-DNA shows no nopaline accumulation. In fact, there is overall a stronger indication of a *negative* correlation between copy number and level of expression than of a positive correlation. However, one would still anticipate that on average there would be a positive correlation between the number of independent loci at which T-DNA has inserted and the level of expression of introduced genes. The experiments presented here do not directly address the question of locus number.

Secondly, only a small proportion of the null phenotypes of either the *nos* and *cab/ocs* genes can be interpreted in terms of the absence of, or the rearrangement of, DNA encoding the unexpressed gene. Plant 521.2A which is Nos^+ and Ocs^- apparently contains a completely intact copy of the *cab/ocs* fusion gene. Many of the Nos^- plants show intact copies of the nopaline synthase gene. In tomato, one out of two Nos^- plants could be explained by a deletion in the *nos*-coding DNA (Jorgensen et al. 1987). The other example was the plant with five T-DNAs in a complex array referred to above.

Thirdly, the inverted repeat structure does not preclude high level expression of the *cab/ocs* gene. This is most clearly illustrated by 571.7B which expresses the *cab/ocs* fusion at a high level. Conversely, the presence of an inverted repeat does not assure high level gene expression as is illustrated by 571.5B. In contrast to the *cab/ocs* chimeric gene, the nopaline synthase gene is only expressed in T-DNA inserts which show some right border sequences in a non-inverted repeat conformation. However, the *nos* gene is only expressed in 6/28 transformants and therefore this is not significant. Also, Jorgensen et al. (1987) have presented data showing that in tomato an IR about the right border does not preclude high level *nos* gene expression.

A fourth conclusion is based on the levels of expression in shoots regenerated from the same callus. These shoots are usually (503.5, 503.9, 503.10, 571.7, 571.9, 571.10) but not always (503.1, 511.2) genotypically identical. Inspection of the levels of expression reported for these plants in Jones et al. (1985) enables us to evaluate the extent to which all the variability in levels of gene expression is due to position-effect variation. There are several examples of genotypically identical individuals that show significantly different levels of expression, especially in the shoots regenerated from calli 503.5 and 503.10. Some of this variability must either be due to some epigenetic differences or due to insufficient physiological standardization of the greenhouse-grown material which was analyzed for expression. We have confirmed that the differences between 503.5A, B, C and D and between 503.10A and 503.10B are not due to innaccurate measurement of the RNA amounts. In 503.10A and 10B the amount of *nos* gene mRNA varies by the same amount as the *cab/ocs* gene mRNA. This observation could be explained if the proportion of polyadenylated mRNA in the total RNA population was very different. The variability between the genotypically identical clones from 571.7, 571.9 and 571.10 is less dramatic than the variability in the plants from 503.5 and 503.10.

We now recognize several possible sources of physiological variability in the experiments of Jones et al. (1985). The amount of mRNA for *cab* genes may fluctuate diurnally (Kloppsteck 1985) and we have since observed this phenomenon in both *nos* and *cab/ocs* genes (J. Jones and D. Gilbert, unpublished). Therefore we now regard it as essential to harvest plant material at the same time of day. Different plant samples when harvested may contain different proportions of young meristematic leaves and older fully expanded leaves, and it is possible that these tissues differ in their mRNA/ribosomal RNA ratio and in their mRNA population. Plants arise from tissue culture at different times and arrive in the greenhouse on different dates and so even though the harvested plants may be of apparently identical appearance they are not in an identical environment.

These data have not led us to withdraw our conclusion about the existence of between-transformant variability. For example, 503.9A, 9B and 9C consistently show zero expression of the introduced genes despite the presence of coding sequences in multiple copies.

This variability phenomenon has been observed by other researchers working with introduced genes (e.g., Nagy et al. 1985) although soybean storage protein genes when assayed in the developing seeds of transgenic petunia and tobacco show lower levels of variability (Beachy et al. 1985; R.B. Goldberg, personal communication). Conceivably, it is easier to standardize conditions for the expression of genes which are predominantly developmentally regulated than for the expression of photosynthetic genes which are also strongly regulated by environmental conditions. Since the gene expression assay involves analysis of accumulated mRNA, mRNA species which may be more stable such as storage protein gene mRNAs (Walling et al. 1986) may show less between-individual fluctuations in mRNA levels and thus display less apparent between-transformant variability.

This detailed analysis of the structures of the introduced T-DNAs has proved useful in trying to interpret the various expression levels which have been observed in individual plants. The analysis has permitted the unambiguous assessment of whether or not individual plants derived from the same callus are somatic sibs and thus an evaluation of the extent to which between-transformant variation can be explained by epigenetic or physiological differences between individuals. These data rule out the idea that quantitative differences in copy number are the main source of quantitative differences in the level of introduced gene expression. We show that in many plants in which an introduced gene is present in an unrearranged state, the expression of this gene cannot be detected. This result makes more attractive the hypothesis that the particular position of insertion of the T-DNA can influence the expression of the introduced gene. However, the relative extent to which between-transformant variation in gene expression is due to genomic position effects, to the effects of the particular structure of the introduced T-DNA, and to the effects of other physiological or epigenetic factors still needs careful investigation.

Acknowledgements. We would like to thank Kathy Lee for making available pKL133, J. Bedbrook and P. Dunsmuir for helpful discussions, R. B. Goldberg for discussing unpublished data, Tom Lemieux for keeping the petunias alive for two years, H. Dooner, G. Warren, C. Dean and P. Dunsmuir for useful criticism of the manuscript and Carol Rubinstein, Connie Stephens and Rob Narberes for assistance in preparing figures.

References

Beachy RN, Chen Z-L, Horsch RB, Rogers SG, Hoffmann NJ, Fraley RT (1985) Accumulation and assembly of soybean beta-

conglycinin in seeds of transformed petunia plants. EMBO J 4:3047–3054
Depicker A, Stachel S, Dhaese P, Zambryski P, Goodman H (1982) Nopaline synthase: transcript mapping and DNA sequence. J Mol Appl Genet 1:561–573
Jones JDG, Dunsmuir P, Bedbrook JR (1985) High level expression of introduced chimaeric genes in regenerated transformed plants. EMBO J 4:2411–2418
Joos H, Inze D, Caplan A, Sormann M, Van Montagu M, Schell J (1983) Genetic analysis of T-DNA transcripts in nopaline crown galls. Cell 32:1057–1067
Jorgensen R, Snyder C, Jones JDG (1987) T-DNA is organised predominately in inverted repeat structures in plants transformed with *Agrobacterium tumefaciens* C58 derivatives. Mol Gen Genet 207:471–477
Kloppsteck K (1985) Diurnal and circadian rhythmicity in the expression of light-induced plant nuclear messenger RNAs. Planta 165:502–506
Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbour Laboratory, Cold Spring Harbour, New York
Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8:4321–4325
Nagy F, Morelli G, Fraley RT, Rogers SG, Chua N-H (1985) Photoregulated expression of a pea *rbcS* gene in leaves of transgenic plants. EMBO J 4:3063–3068
Southern EM (1975) Detection of specific sequences among DNA fragments separated by electrophoresis. J Mol Biol 98:503–517
Walling L, Drews GN, Goldberg RB (1986) Transcriptional and post-transcriptional regulation of soybean seed protein mRNA levels. Proc Natl Acad Sci USA 83:2123–2127
Zambryski P, Joos H, Genetello C, Leemans J, Van Montagu M, Schell J (1983) Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. EMBO J 2:2143–2150

Communicated by R.B. Goldberg

Received December 3, 1986