Volume 17 Number 7 1989 Nucleic Acids Research

Stability and expression of bacterial genes in replicating geminivirus vectors in plants

R.J.Hayes, R.H.A.Coutts and K.W.Buck*

Department of Pure and Applied Biology, Imperial College of Science, Technology and Medicine, London SW7 2BB, UK

Received February 7, 1989; Revised and Accepted March 8, 1989

ABSTRACT

Bacterial beta-glucuronidase (*gus*) and neomycin phosphotransferase (*neo*) genes were introduced into coat protein replacement vectors based on DNA A of tomato golden mosaic virus (TGMV). Recombinant *gus* and *neo* vectors up to 1.1 kbp larger than DNA A were shown to replicate stably in transgenic plants containing partial dimers (master copies) of the vectors integrated into their chromosomal DNA in the absence of DNA B. Beta-glucuronidase and neomycin phosphotransferase activities in independently transformed plants were proportional to the copy number of the double-stranded forms of the vector. Deletion analysis has shown that an essential part of the TGMV coat protein promoter, including a TATA box, lies within 76 nt upstream of the initiation codon of the gene. An increase in expression of a *neo* gene was obtained by replacing this 76 nt sequence by an 800 nt sequence containing a cauliflower mosaic virus 35S RNA promoter with no effect on the ability of the vector to replicate or on its stability in transgenic plants. Systemic infection of plants by agroinoculation with TGMV vectors larger than DNA A in the presence of DNA B resulted in deletions in the vector DNA in some, but not all, plants. Possible reasons for vector instability in systemically infected plants, and vector stability in transgenic plants containing master copies of the vector, are discussed.

INTRODUCTION

Tomato golden mosaic geminivirus (TGMV) has a genome of two circular single-stranded (ss) DNA molecules, DNA A (2588 nucleotides) and DNA B (2508 nucleotides)(1). DNA A encodes proteins required for virus DNA replication (2,3) and can replicate in single cells in the absence of DNA B (4, 5). DNA B, which is dependent on DNA A for its replication, is required for the virus to spread from cell-to-cell leading to systemic infection of the plants (4, 6). The virus coat protein, encoded by open reading frame (ORF) AR1 of DNA A (7), is not required for either replication in single cells or systemic spread of viral DNA through plants (2, 8, 9)

Recently we constructed a vector, TA*neo*, in which the coat protein coding region of DNA A was replaced by a bacterial neomycin phosphotransferase gene (*neo* gene)(10). The vector was introduced into plants in two ways. In the first, transgenic plants containing a partial dimer of TA*neo* integrated into their chromosomal DNA were produced *via* agrobacterium-mediated transformation. In such plants monomeric TA*neo* DNA was produced, either by a replicative mechanism or by intramolecular recombination, and replicated. DNA B was not required because every cell contained a master copy of the vector integrated into its chromosomal DNA. In the second method, a partial dimer of the vector was introduced into the stem base of a transgenic plant, containing a dimer of DNA B integrated into its chromosomal DNA (B2 plant) by agrobacterium-mediated

inoculation (agroinoculation)(8, 11). In this procedure, formation and replication of monomeric TA*neo* in a single cell can lead to formation and replication of monomeric DNA B, leading to cell-to-cell transmission and systemic infection of the plant with TA*neo*. With both methods replication of the vector resulted in amplification of the *neo* gene and of the level of neomycin phosphotransferase activity (NPT activity).

Apart from geminivirus vectors, the only other DNA vectors that have been used for gene amplification in plants are based on cauliflower mosaic virus (CaMV) (12). Two potential advantages of geminivirus vectors are as follows. (i) Because the coat protein gene has been deleted, the size of the DNA inserted will not be limited by packaging constraints, as it is with CaMV vectors. However there may be other constraints that could affect the replication and stability of vectors larger than DNA A. Recently replication of a TGMV vector larger than DNA A was demonstrated using a transient expression system in leaf discs (13), but there is no information on the replication and stability of vectors larger than DNA A in transgenic plants containing chromosomal master copies of the vector or in agroinoculated plants. (ii) It may be possible to replace geminivirus coat protein promoters, allowing studies of the transcription of inserted genes from other promoters. In CaMV vectors the promoter controlling expression of inserted genes also controls expression of essential virus genes and formation of an RNA intermediate of CaMV DNA replication (12). However it remains to be demonstrated that removal of the coat protein promoter, and insertion of a heterologous promoter, has no effect on the replication or stability of a geminivirus vector. In the present paper we show that TGMV vectors at least up to 1.1 kbp larger than the native DNA A can be replicated stably, and inserted genes amplified and expressed, in transgenic plants containing a partial dimer of the vector integrated into the plant chromosomal DNA in the absence of DNA B. An increase in expression of an inserted gene was obtained by replacing a 76 nt sequence upstream of the TGMV coat protein coding region by an 800 nt sequence containing a CaMV 35S RNA promoter with no effect on the ability of the vector to replicate or on its stability in transgenic plants. However vectors larger than native DNA A were unstable in plants systemically infected by agroinoculation with the vector in the presence of DNA B. In a proportion of such plants, deletions in the vector occurred which resulted in the generation of DNA molecules with sizes down to that of DNA A.

METHODS

Molecular biological procedures

The following procedures were carried out as described previously (2, 8, 10): enzyme incubations; gel electrophoresis; Southern blotting; recombinant DNA manipulations; agroinoculation of transgenic B2 *Nicotiana tabacum* cv. Samsun plants; agrobacterium-mediated transformation of *N. tabacum* cv. Samsun plants; production of transgenic pBin19, TA*neo*, A1.6*neo* and B2 plants; determination of copy numbers of double-stranded forms of vector molecules; measurement of NPT activity.

Plasmids

pTAgus. Plasmid pAE (TGMV DNA A cloned into the *Eco*RI site of pEMBL9)(2) was cleaved partially with *Xho*II and then with *Asu*II. The resultant 1.9 kbp fragment, which lacks most of the coat protein gene including the AUG initiation codon but retains the coat protein promoter, was separated by gel electrophoresis, made blunt-ended with T4 DNA polymerase and recircularised after addition of *Bgl*II linkers to give pABg. Plasmid pBI221, which consists of the 3.0 kbp *Hin*dIII–*Eco*RI fragment of pBI121 (ref. 14)

containing the CaMV 35S RNA promoter-beta-glucuronidase-*nos* poly A site cloned between the corresponding sites of pUC19, was cut with *Sst*I, made blunt-ended with mung bean nuclease and recircularised after addition of *Bgl*II linkers. The product was cut with *Bgl*II and *Bam*H1 and the 1.8 kbp fragment, containing the beta-glucuronidase (*gus*) gene but not the CaMV 35S RNA promoter or *nos* poly A site, was cloned into the *Bgl*II site of pABg to give pTA*gus*.

pTAPLneo. Plasmid pBH401 (TGMV DNA A cloned into the *Eco*RI site of pAT153)(14) was cleaved with *Hin*dIII and *Dra*I and the 0.6 kbp fragment between the *Hin*dIII site of pAT153 and the *Dra*I site at nucleotide 259 of DNA A was isolated by gel electrophoresis. pNEO (Pharmacia) was cut with *Bgl*II and the linearised plasmid was end-filled using Klenow polymerase. After cleavage with *Hin*dIII the 5.4 kbp fragment was ligated to the 0.6 kbp *Hin*dIII-*Dra*I fragment above to form a circular molecule. This was cleaved with *Hin*dIII and *Asu*II and the 1.4 kbp fragment was then cloned between the *Hin*dIII and *Asu*II sites of pBH401 to generate pTAPL*neo*.

pTA35Sneo and pTAS53neo. pBI221 was cleaved with *Hin*dIII and the product was end-filled, followed by addition of a *Bgl*II linker. After cleavage with *Bam*H1 and *Bgl*II, the 0.8 kbp fragment containing the CaMV 35S RNA promoter was cloned into the *Bcl*I site of pTAPL*neo* in the forward and reverse orientations to give pTA35S*neo* and pTAS53*neo* respectively.

Plasmids pA1.6gus, pA1.6PLneo, pA1.6−35Sneo and pA1.6−S53neo were constructed by cleaving pTA*gus*, pTAPL*neo*, pTA35S*neo* and pTAS53*neo* respectively with *Eco*RI, separation of the inserts TA*gus*, TAPL*neo*, TA35S*neo* and TAS53*neo* from the plasmid vector moiety by gel electrophoresis, and cloning each insert into the *Eco*RI site of pA0.6 (a 1.7 kbp fragment of DNA A cloned into the polylinker region of a modified pEMBL9 vector, ref. 8).

Plasmids pBin19A1.6, pBin19A1.6gus, pBin19A1.6PLneo, pBin19A1.6−35Sneo, pBin19A1.6−S53neo, pBin19TAgus, pBin19TA35Sneo and pBin19−35Sgus were constructed by cleaving pA1.6, pA1.6*gus*, pA1.6PL*neo*, pA1.6−35S*neo*, pA1.6−S53*neo*, pTA*gus*, pTA35S*neo* and pBI221 with *Hin*dIII (for each of which there is one site in the plasmid vector moiety, but no site in DNA A, *neo*, *gus* or the CaMV 35S RNA promoter) and cloning each linearised DNA into the *Hin*dIII site of pBin19 (ref. 16).

Plasmids pTA*neo*3, pA1.6*neo* and pBin19A1.6*neo* have been described previously (8, 10).

Preparation of DNA

DNA from agroinfected or transgenic plants was prepared by the method of Hamilton et al. (17) when the DNA was to be analysed by gel electrophoresis and Southern blotting, and by the method of Lichtenstein and Draper (18) for determination of vector copy number.

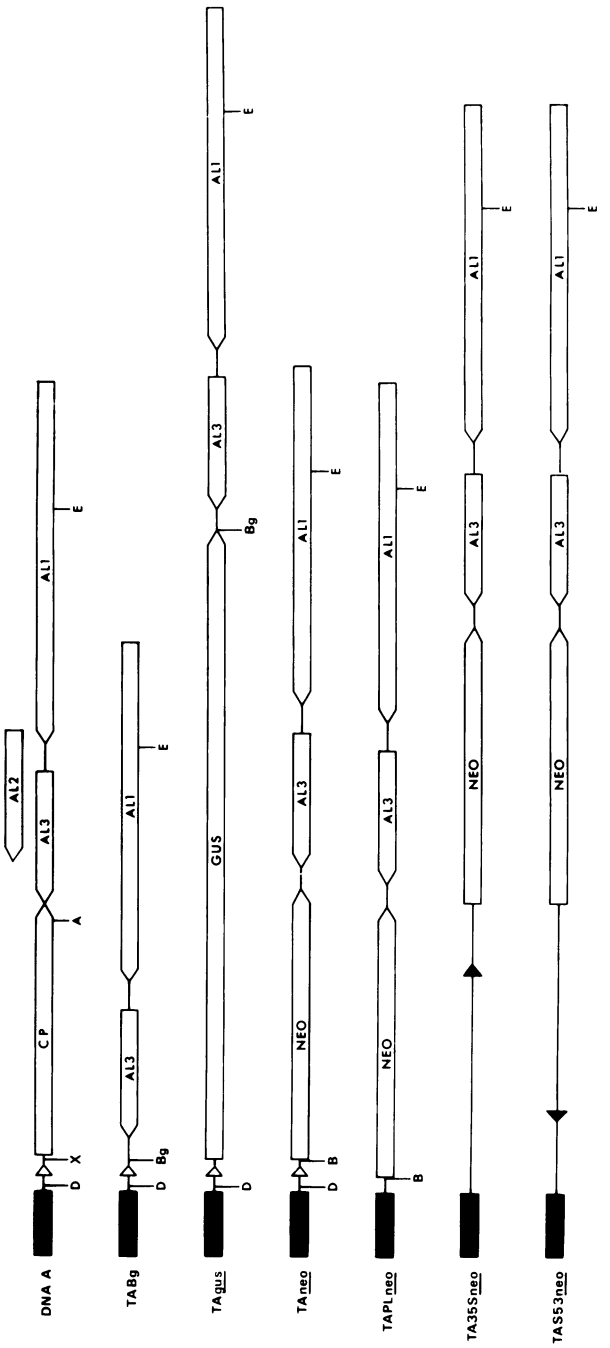
GUS activity

This was measured as described by Jefferson et al. (14).

RESULTS

Replicating TGMV vectors larger than DNA A are stable in transgenic plants containing chromosomal master copies of the vector.

A beta-glucuronidase gene linked to a TGMV coat protein promoter can be stably amplified and expressed in transgenic plants. Previously we showed that a *neo* gene could be stably replicated by a TGMV vector both in transgenic plants containing a master copy of the vector and in plants agroinoculated with the vector (10). The size of the *neo* gene (0.79 kbp) is similar to that of the coat protein gene (0.74 kbp) and the size of TA*neo* (2.7 kbp)


DNA A
AL2
AL3
AL1
E
CP
A
D
X
TABg
AL1
E
AL3
Bg
D
TAgus
AL1
E
AL3
Bg
GUS
D
TAneo
AL1
E
AL3
NEO
D
B
TAPLneo
AL1
E
AL3
NEO
B
TA35Sneo
AL1
E
AL3
NEO
TAS53neo
AL1
E
AL3
NEO

is similar to that of native DNA A (2.6 kbp). In order to determine if genes significantly larger than the coat protein gene could be stably replicated and expressed, we selected a bacterial beta-glucuronidase gene (*gus* gene, 1.8 kbp) (19) for study.

A TGMV vector, TABg, was constructed in which the region between the *Xho*II site (nt 311) and the *Asu*II site (nt 1037) of DNA A (containing most of the coat protein coding region, including the AUG initiation codon at nt 327) was deleted and a *Bgl*II site inserted (Fig. 1). This vector should enable any gene cloned into the *Bgl*II site to be expressed in plants under the control of the TGMV coat protein promoter (tentatively associated with a TATA box starting at nt 287/289; ref. 1, 20). The *gus* gene was cloned into the *Bgl*II site of TABg to give TA*gus* (Fig. 1).

To produce transgenic plants in which replicating TA*gus* molecules could be produced from a master copy of the vector inserted into the plant chromosomal DNA, TA*gus* was cloned into the *Eco*RI site of pA0.6. The resultant plasmid pA1.6*gus* contains two copies of open reading frame AL1 and the 200bp 'common region' (thought to contain an origin of DNA replication, ref. 1) to allow release of TA*gus* in plants by recombination or replication. pA1.6*gus* was cloned into the Ti plasmid vector pBin19 to give pBin19A1.6*gus* and transgenic *N. tabacum* plants containing chromosomal insertions of A1.6*gus* (transgenic A1.6*gus* plants) were produced using agrobacterium-mediated transformations. For comparison transgenic plants containing chromosomal insertions of a partial dimer of DNA A (transgenic A1.6 plants), a monomer of TA*gus* (transgenic TA*gus* plants) and the CaMV 35S RNA promoter-*gus-nos* polyA site chimaera (transgenic 35S*gus* plants) were also produced.

Gel electrophoresis and Southern blotting of DNA from transgenic A1.6*gus* plants revealed ssDNA and supercoiled and open-circular double-stranded (ds) DNA forms corresponding in size to the monomeric TA*gus* (3.7kbp)(Fig. 2, lane 3). These migrated more slowly than the corresponding forms of DNA A detected in transgenic A1.6 plants (Fig. 2, lane 1). The copy number of the dsDNA forms of TA*gus* ranged from 85 to 140 in 10 independently transformed plants and the beta-glucuronidase activity (GUS activity) in plant extracts paralleled the copy number (Table 1). There was no evidence for deletions to generate molecules smaller than TA*gus* in any of the transgenic A1.6*gus* plants (Fig. 2, lanes 5 to 14). No freely replicating TA*gus* DNA was detected in transgenic TA*gus* plants which contain monomeric copies of TA*gus* inserted into their chromosomal DNA and the GUS activities of these plants (and of the transgenic 35S*gus* plants) were correspondingly low (Table 1).

Deletion of a 76 nt sequence upstream of the coat protein coding region prevents gene expression but has no effect on vector replication or stability. A TGMV vector, TAPL*neo*, was constructed in which the region between the *Dra*I site (nt 259) and *Asu*II site (nt 1037) of DNA A had been deleted and a promoter-less *neo* gene inserted (Fig. 1). TAPL*neo* lacks most of the coat protein gene and 76 nucleotides upstream of the initiation codon, including a TATA box tentatively associated with the coat protein promoter. Transgenic plants containing a partial dimer of TAPL*neo* inserted into their chromosomal DNA

Figure 1. Structures of TGMV DNA A and derived vectors. The circular vectors are shown in linear form starting from nucleotide 1 of the DNA A sequence (1). The black box represents the 200 bp 'common region'. Open triangle, TGMV coat protein promoter. Closed triangle, CaMV 35S RNA promoter. CP, coat protein coding region. GUS, beta-glucuronidase coding region. NEO, neomycin phosphotransferase coding region. AL1, AL2 and AL3, open reading frames in TGMV DNA A; AL2 is present in all the vectors but is shown only for DNA A. A = *Asu*II, B = *Bcl*I, Bg = *Bgl*II, D = *Dra*I, E = *Eco*RI, X = *Xho*II.

Table 1. Gene copy numbers and GUS activity

Plant		Average gene copy number per *N. tabacum* genome	GUS activity[a]	GUS activity per unit gene copy number
Transgenic pBin19[b]		7	None detected	
Transgenic TA*gus*[b]		4	0.19	
Transgenic 35S*gus*[b]		4	0.40	
Transgenic A1.6*gus*	(1)[c]	85	3.6	0.042
	(2)	90	3.8	0.042
	(3)	95	3.9	0.044
	(4)	95	4.2	0.041
	(5)	100	3.6	0.036
	(6)	100	4.0	0.040
	(7)	110	5.0	0.045
	(8)	120	5.2	0.043
	(9)	135	5.8	0.043
	(10)	140	6.2	0.044
Transgenic B2, agroinfected with A1.6*gus*[d]	(1)	85	3.5	0.041
	(2)	120	5.5	0.046
	(3)	280	11	0.039
	(4)	370	16	0.043
	(5)	400	17	0.043
	(6)	450	19	0.042

[a] Moles 4-methyl umbelliferone/min/mg protein
[b] Highest values obtained from a range of independent transformants
[c] Values for ten individual plants, independently transformed
[d] Values for six individual plants

(transgenic A1.6PL*neo* plants) were constructed in a similar way to transgenic A1.6*gus* plants. Southern blots of DNA from such plants revealed replicating monomeric ssDNA and dsDNA forms of TAPL*neo* (2.6 kbp) (Fig. 3). The TAPL*neo* copy number in eight

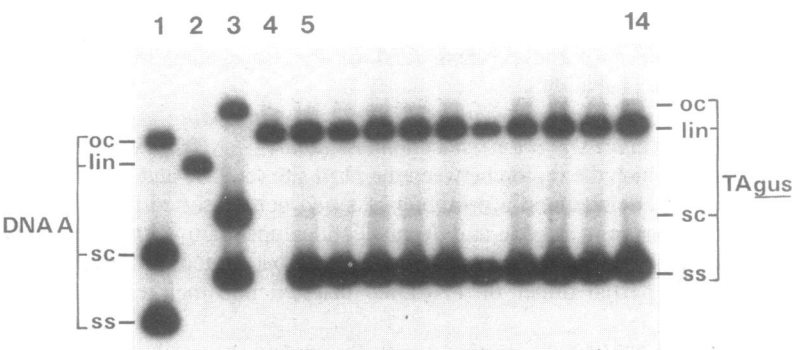


Figure 2. Southern blot analysis of DNA from transgenic A1.6*gus* plants. Lane 1, DNA from transgenic A1.6 plant; lane 2, *Eco*RI-digested pBH401; lane 3, DNA from transgenic A1.6*gus* plant; lane 4, *Eco*RI-digested pTA*gus*; lanes 5 to 14, *Eco*RI-digested DNA from 10 different transgenic A1.6*gus* plants. The blot was hybridised with a ^{32}P-labelled DNA A-specific probe. The positions of the single-stranded (ss) DNA and open circular (oc), linear (lin) and supercoiled (sc) double-stranded forms of DNA A and TA*gus* are shown on the side of the gel.

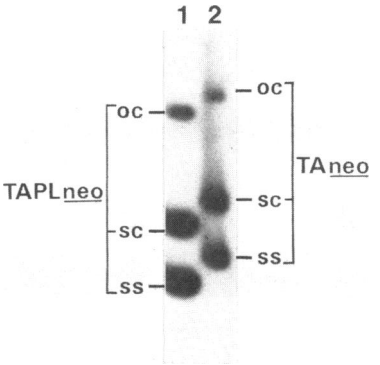


Figure 3. Southern blot analysis of DNA from transgenic A1.6PL*neo* plant. Lane 1, DNA from transgenic A1.6PL*neo* plant; lane 2, DNA from transgenic A1.6*neo* plant. The probe and abbreviations are as in Figure 2. The positions of the ssDNA and dsDNA forms of TAPL*neo* and TA*neo* are shown on the sides of the gel.

independently transformed plants ranged from 80−120, but only very low levels of NPT activity, similar to those found in transgenic pBin19 plants (Table 2), could be detected in extracts of any of the plants (results not shown).

Insertion of a CaMV 35S RNA promoter increases gene expression with no effect on vector replication or stability. A DNA fragment containing the CaMV 35S RNA promoter was cloned into the *Bcl*I site of TAPL*neo*, upstream of the *neo* gene, in the forward and reverse orientations to create TA35S*neo* and TAS53*neo* respectively (Fig. 1). Transgenic plants containing a partial dimer of TA35S*neo* (transgenic A1.6−35S*neo* plants), a partial dimer of TAS53*neo* (transgenic A1.6-S53*neo* plants) or monomer of TA35S*neo* (transgenic TA35S*neo* plants), were constructed. Southern blots of DNA from nine independently transformed A1.6−35S*neo* plants revealed monomeric forms of ssDNA and dsDNA corresponding in size to TA35S*neo* (3.4 kbp)(Fig. 4) with no evidence of deletions to form smaller molecules. Determination of copy numbers and NPT activities in these plants showed

Table 2. Gene copy numbers and NPT activity

Plant	Average gene copy number per *N tabacum* genome	Relative NPT activity	Relative NPT activity per unit gene copy number
Transgenic pBin19[a]	7	1	
Transgenic TA*neo*[a]	7	3.5	
Transgenic TA35S*neo*[a]	4	5	
Transgenic A1.6*neo*[b]	120	60	0.50
Transgenic A1.6−35*neo*[b]	100	120	1.20
Transgenic B2, agroinoculated with A1.6*neo*[c]	490	240	0.49
Transgenic B2, agroinoculated with A1.6−35S*neo*[c]	450	520	1.15

[a] Highest values obtained from a range of independent transformants.
[b] Average values from nine individual plants, independently transformed.
[c] Average values from ten individual agroinoculated plants.
[b,c] All values from individual plants were within 20% of the average value.

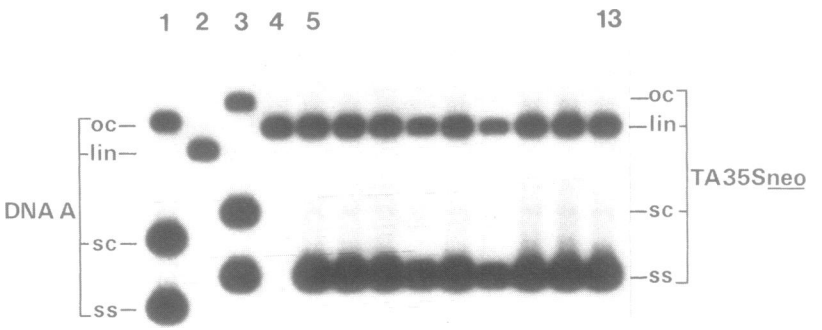


Figure 4. Southern blot analysis of DNA from transgenic A1.6.35S*neo* plants. Lane 1, DNA from transgenic A1.6 plant; lane 2, *Eco*RI-digested pBH401; lane 3, DNA from transgenic A1.6−35S*neo* plant; lane 4, *Eco*RI-digested pTA35S*neo*; lanes 5 to 13, *Eco*RI-digested DNA from 9 different transgenic A1.6−35S*neo* plants. The probe and abbreviations are as in Figure 2. The positions of the ssDNA and dsDNA forms of DNA A and TA35S*neo* are shown on the sides of the gel.

that both the *neo* gene and NPT activity had been amplified, compared with transgenic plants containing only chromosomal monomeric copies of the *neo* gene (transgenic TA35S*neo* plants)(Table 2). Furthermore the NPT activity per unit ds copy number of the *neo* gene was higher in plants in which the *neo* gene was linked to the CaMV 35S RNA promoter (transgenic A1.6−35S*neo* plants) compared with plants in which the *neo* gene was linked to the TGMV coat protein promoter (transgenic A1.6*neo* plants). In transgenic A1.6-S53*neo* plants, the monomeric vector TAS53*neo* (3.4 kbp) was formed (Fig. 5) and replicated to a copy number similar to that of TA35S*neo* in transgenic A1.6−35S*neo* plants. However only very low levels of NPT activity, similar to those in extracts of transgenic pBin19 plants (Table 2), were detected in any of these plants.

Deletions can occur in TGMV vectors larger than DNA A during replication and spread in agroinoculated plants

When 40 transgenic B2 *N. tabacum* plants were agroinoculated with pBin19A1.6*gus*, 21 plants became infected, as determined by dot hybridisation of DNA from non-inoculated leaves with a *gus* gene probe 21 days later. Gel electrophoresis and Southern blotting of

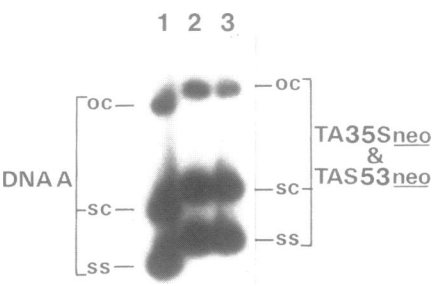


Figure 5. Comparison of the DNA forms in transgenic A1.6−35S*neo* and A1.6-S53*neo* plants by Southern blotting. Lane 1, DNA from transgenic A1.6 plant; lane 2, DNA from transgenic A1.6−35S*neo* plant; lane 3, DNA from transgenic A1.6-S53*neo* plant. The probe and abbreviations are as in Figure 2. The positions of the ssDNA and dsDNA forms of DNA A, TA35S*neo* and TAS53*neo* are shown on the sides of the gel.

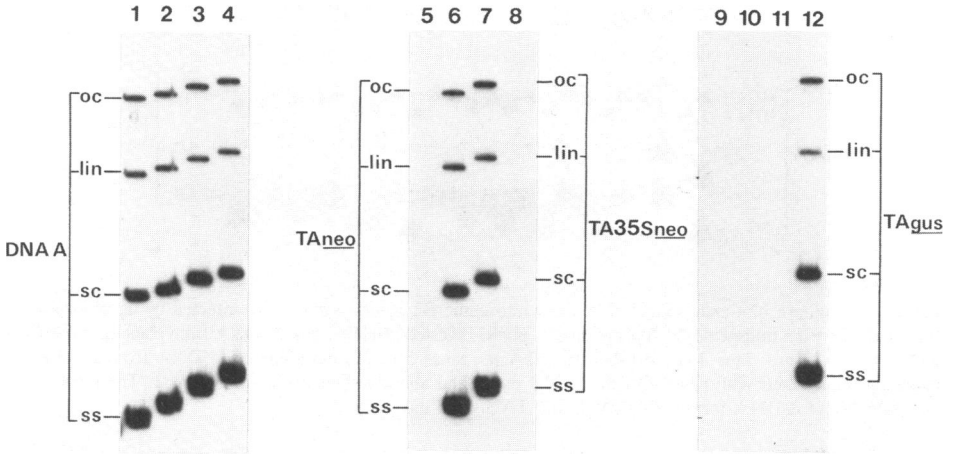


Figure 6. Southern blot analysis of DNA from transgenic B2 plants systemically infected by agroinoculation. Lanes 1, 5 and 9, DNA from plants agroinoculated with pBin19A1.6; lanes 2, 6 and 10, DNA from plants agroinoculated with pBin19A1.6*neo*; lanes 3, 7 and 11, DNA from plants agroinoculated with pBin19A1.6−35S*neo*; lanes 4, 8 and 12, DNA from plants agroinoculated with pBin19A1.6*gus*. The probes were: lanes 1 to 4, as in Figure 2; lanes 5 to 8, ^{32}P-labelled *neo* gene; lanes 9 to 12, ^{32}P-labelled *gus* gene. The abbreviations are as in Figure 2. The positions of the ssDNA and dsDNA forms of DNA A, TA*neo*, TA35S*neo* and TA*gus* are indicated.

DNA from six of the infected plants revealed ssDNA and supercoiled, open circular and linear dsDNA forms corresponding in size to full-length monomeric TA*gus* (3.7 kbp), similar to those detected in transgenic A1.6*gus* plants (Fig. 2), using either a TGMV DNA A probe or a *gus* probe (Fig. 6, lanes 4, 8 and 12; Fig. 7, lanes 3, 5, 8, 15, 17, 20 and 23). Copy numbers of the dsDNA forms of the TA*gus* vector were variable, ranging from 85 to 450 in different plants. However there was a good correlation between the level of GUS activity and the vector copy number (Table 1).

Southern blotting of DNA from non-inoculated leaves of the other 15 infected plants showed that the TA*gus* vector had undergone deletions resulting in molecules of a range of sizes down to that of native DNA A (2.6 kbp)(Fig. 7). No GUS activity could be detected in extracts of any of these plants.

When 40 transgenic B2 *N. tabacum* plants were agroinoculated with pBin19A1.6−35S*neo*, 29 plants became infected as determined by hydridisation of DNA from non-inoculated leaves to a CaMV 35S RNA promoter probe 21 days later. Southern blotting of DNA from 16 of the infected plants revealed ssDNA and supercoiled, open-circular and linear dsDNA forms corresponding in size to TA35S*neo* (3.4 kbp), similar to those detected in transgenic A1.6−35S*neo* plants, with either DNA A or *neo* probes. (Fig. 6, lanes 3 and 7; Fig. 8). Copy numbers of the ds forms of TA35S*neo*, measured in 10 of these plants, were less variable than in plants agroinoculated with pBin19A1.6*gus* (Table 1). Both vector copy number and NPT activity were higher than those in transgenic A1.6−35S*neo* plants (Table 2). However NPT activities in plants agroinoculated with pBin19A1.6−35S*neo* were higher than those in plants agroinoculated with pBin19A1.6*neo*, even though vector copy numbers were similar.

Southern blotting of DNA from non-inoculated leaves of the remaining 13 infected plants showed that the TA35S*neo* vector had undergone deletions resulting in molecules with

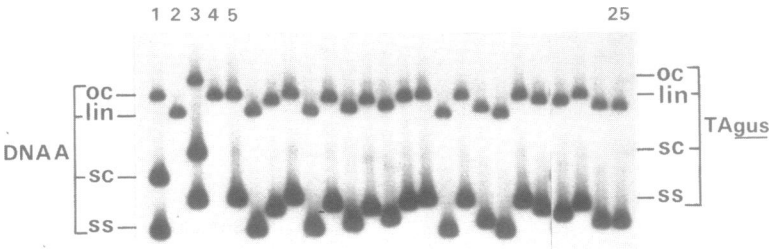


Figure 7. Southern blot analysis of DNA from transgenic B2 plants systemically infected by agroinoculation. Lane 1, DNA from transgenic A1.6 plant; lane 2, *Eco*RI-digested pBH401; lane 3, DNA from plant agroinoculated with pBin19A1.6*gus*; lane 4, *Eco*RI-digested pTA*gus*; lanes 5 to 25, *Eco*RI-digested DNA from 21 different plants agroinoculated with pBin19A1.6*gus*. The probe and abbreviations are as in Figure 2. The positions of the ssDNA and dsDNA forms of DNA A and TA*gus* are indicated.

sizes down to that of native DNA A. Ten of these are shown in Fig. 8. No NPT activity could be detected in extracts of any of these plants.

DISCUSSION

We have shown that TGMV vector molecules, at least up to 1.1 kbp larger than native viral DNA A, can be stably replicated in transgenic plants containing chromosomal master copies (partial dimers) of the vector. Hence the coat protein replacement vector TABg (Fig. 1) should be suitable for stable amplification of genes up to at least 1.8 kbp and their expression under the control of the TGMV coat protein promoter. Failure to detect any deletions in the largest vector studied (3.7 kbp) in transgenic plants suggests that even larger genes could probably be replicated stably. Since the ability of monomeric DNA A molecules to form and replicate in transgenic A2 plants is inherited as a Mendelian trait (4), amplification of genes up to at least 1.8 kbp should be both stable and heritable.

Deletion of 76 nucleotides upstream of the TGMV coat protein initiation codon had no effect on DNA replication, since vector copy numbers in transgenic A1.6*neo* and A1.6PL*neo* plants were similar. However whereas NPT activity in transgenic A1.6*neo* plants was amplified 60 fold over the background activity due to expression of the chromosomal *neo* gene in the pBin19 vector, there was no amplification of NPT activity

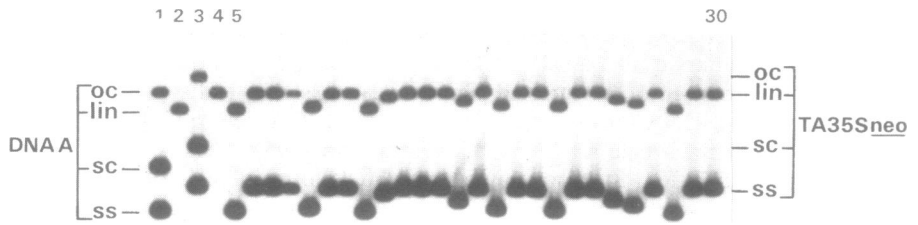


Figure 8. Southern blot analysis of DNA from transgenic B2 plants sytemically infected by agroinoculation. Lane 1, DNA from transgenic A1.6 plant; lane 2, *Eco*RI-digested pBH401; lane 3, DNA from plant agroinoculated with pBin19A1.6−35S*neo*; lane 4, *Eco*RI-digested pTA35S*neo*; lanes 5 to 30, *Eco*RI-digested DNA from 26 different plants agroinoculated with pBin19A1.6−35S*neo*. The probe and abbreviations are as in Figure 2. The positions of the ssDNA and dsDNA forms of DNA A and TA35S*neo* are indicated.

in transgenic A1.6PL*neo* plants. This indicates that an essential part of the coat protein promoter lies within 76 nucleotides of the initiation codon. Analysis of the nucleotide sequence of DNA A (1) revealed potential promoter sequences, corresponding to the consensus sequence TATAT/AA (21), starting at nt 89 and 287/289 (overlapping). The position of the 5′ terminus of a transcript spanning the coat protein coding region indicated that the TATA box at nt 287/289 is the one most likely to be associated with the coat protein promoter (20). The evidence presented here is consistent with this conclusion.

Insertion of the CaMV 35S RNA promoter into TAPL*neo* in either orientation had no effect on DNA replication, since vector copy numbers in transgenic A1.6PL*neo*, A1.6−35S*neo* and A1.6-S53*neo* plants were similar. However amplification of NPT activity was observed only when the 35S promoter was in the correct orientation. The levels of expression per unit vector copy number of the *neo* gene linked to the CaMV 35S RNA promoter (in transgenic A1.6−35S*neo* plants and in plants agroinoculated with A1.6−35S*neo*) were more than double those in which the *neo* gene was linked to the TGMV coat protein promoter (in transgenic A1.6*neo* plants and in plants agroinoculated with A1.6*neo*). This shows that the promoter region in TGMV coat protein replacement vectors can be manipulated to obtain higher levels of gene expression.

The increase in levels of gene expression obtained by inserting a CaMV 35S RNA promoter into TAPL*neo* suggests that TGMV vectors have potential for development as promoter expression vectors. Ti plasmid vectors containing promoter-less reporter genes have been used to identify sequences with plant promoter activity (22). TGMV vectors could have two advantages over Ti vectors for this purpose. (i) Since TGMV vectors replicate to high copy number, assays would be much more sensitive. (ii) The expression of genes inserted into plant chromosomes can vary substantially (up to 200 fold) in independently transformed plants and there is often no correlation between gene copy number and level of expression (22−25). One possible explanation is that expression depends on the position of insertion within a chromosome (positional effects). Expression from the extrachromosomal geminivirus vectors should be independent of positional effects. The levels of GUS activity per unit vector copy number in 10 plants independently transformed with A1.6*gus*, and in 6 plants agroinoculated with A1.6*gus*, were all similar (Table 1). Likewise, levels of NPT activity per unit vector copy number in 8 plants independently transformed with A1.6*neo* and in 10 plants agroinoculated with A1.6 *neo* were similar (Table 2). Hence expression of genes from the TGMV vectors is not subject to positional or other effects which cause variability in expression of chromosomally inserted genes. However before TGMV vectors can be used for analysis of plant promoters, all the *cis* regulatory elements required for maximal expression of the TGMV coat protein promoter must be defined. Although essential sequences of the coat protein promoter including a TATA box, are deleted in TAPL*neo*, there may be other elements, such as enhancers, which contribute to promoter activity. The effect of such sequences on the activity of an inserted promoter, or the extent to which such sequences could be deleted without affecting DNA replication, remain to be evaluated.

No instability of geminivirus vectors with a size similar to that of native viral DNA A has been observed in systemically infected plants, either in the current work or previously (10, 26). However in plants systemically infected with TA*gus* (3.7 kbp) or TA35S*neo* (3.4 kbp), which are 1.1 kbp and 0.8 kbp larger than DNA A respectively, deletions occurred in the vector in a proportion of plants. When deletions did occur, mixtures of deleted and undeleted DNA were never observed. The smaller DNA molecules may

replicate faster, or be transported from cell-to-cell more efficiently, eventually displacing the larger molecules completely.

Deletions in the vector could occur intramolecularly (for a review of possible mechanisms, see ref. 27) or by intermolecular recombination. Stanley and Townsend (28) found that when plants were infected by mechanical inoculation with intact recombinant clones of cassava latent geminivirus DNA in M13 and pUC vectors, intermolecular recombination between the two genome segments occurred generating mutants with sizes similar to those of the native viral DNA components. Although the size selection of mutants was not as stringent in our studies, it is noteworthy that no deleted molecules smaller than DNA A were found with vectors containing either the *gus* or *neo* gene. The size selection of deleted vector molecules could be governed by a competitive replication mechanism with DNA B, as suggested by Lazarowitz (29).

We have not determined the positions of the deletions in any of the smaller molecules generated. However the deletions are likely to be in either the reporter gene or the promoter to which it is attached, since the ability of the deleted vector molecules to replicate and spread through the plants was not impaired. Of the other regions of the vector, ORFs AL1 and AL2 have been shown to be required for infection of plants (2,3) and the 200 bp 'common region', which probably contains one or more origins of DNA replication (1), is also likely to be essential for DNA replication. The absence of GUS activity, and failure to amplify NPT activity, in plants in which deletion mutants were generated, is consistent with a deletion in the reporter gene or its promoter.

The absence of deleted vector molecules in transgenic plants containing master copies of the vector could be due to the absence of replicating DNA B molecules, precluding intermolecular recombination between the two genome segments, and/or the much smaller number of cycles of vector replication. In such transgenic plants, vector molecules replicate only in single cells, with no cell-to-cell transmission. Hence deletion mutants would have much less opportunity to displace a full-length vector molecule than in agroinoculated plants in which the DNA may be introduced into only a single cell and then subsequently undergoes many cycles of replication in different cells as it spreads through the plant.

The highest copy numbers (of intact vector molecules) obtained in leaf tissue from agroinoculated plants were three to four times as great as the highest copy numbers obtained in comparable leaf tissue from transgenic plants containing master copies of the vector (Tables 1 and 2). Whether the vector replicates evenly in all cells within a leaf is not known for either agroinoculated plants or transgenic plants containing master copies of the vector. Hence it is possible that the vector could replicate in a higher proportion of leaf cells in some agroinoculated plants. Alternatively it is possible that DNA B, although not an absolute requirement for DNA replication (4), has a stimulating effect on DNA replication.

In conclusion we have shown that stable amplification and expression of genes up to at least 1.8 kbp can be achieved in plants by integration of master copies of TGMV coat protein replacement vectors into the plant chromosomal DNA. We have also shown that expression of an amplified reporter gene can be increased by replacing sequences of the TGMV coat protein promoter by a heterologous promoter. Finally we have shown that deletions can occur in TGMV vectors larger than DNA during systemic spread through plants in the presence of DNA B.

ACKNOWLEDGEMENTS

We thank the Science and Engineering Research Council Biotechnology Directorate for

the award of a research grant to K.W.B. and R.H.A.C. Experiments were carried out with a licence (no. PHF 29B/116 (107)) issued under the Plant Pests (Great Britain) Order 1980 by the Ministry of Agriculture, Fisheries and Food.

*To whom correspondence should be addressed

REFERENCES

1. Hamilton, W.D.O., Stein, V.E., Coutts, R.H.A. and Buck, K.W. (1984). EMBO J. 3, 2197–2205.
2. Brough, C.L., Hayes, R.J., Morgan, A.J., Coutts, R.H.A. and Buck, K.W. (1988). J. Gen. Virol. 69, 503–514.
3. Elmer, J.S., Brand, L., Sunter, G., Gardiner, W.E., Bisaro, D.M. and Rogers, S.G. (1988). Nucl. Acids Res. 16, 7043–7060.
4. Rogers, S.G., Bisaro, D.M., Horsch, R.B., Fraley, R.T., Hoffmann, N.L., Brand, L., Elmer, J.S. and Lloyd, A.M. (1986). Cell 45, 593–600.
5. Sunter, G., Gardiner, W.E., Rushing, A.E., Rogers, S.G. and Bisaro, D.M. (1987). Plant Mol. Biol. 8, 477–484.
6. Hamilton, W.D.O., Bisaro, D.M., Coutts, R.H.A. and Buck, K.W. (1983). Nucl. Acids Res. 11, 7387–7396.
7. Kallender, H., Petty, I.T.D., Stein, V.E., Panico, M., Blench, I.P., Etienne, A.T., Morris, H.R., Coutts, R.H.A. and Buck, K.W. (1988). J. Gen. Virol. 69, 151–1357.
8. Hayes, R.J., Coutts, R.H.A. and Buck, K.W. (1988). J. Gen. Virol. 69, 1487–1496.
9. Gardiner, W.E., Sunter, G., Brand, L., Elmer, J.S., Rogers, S.G. and Bisaro, D.M. (1988). EMBO J. 7, 899–904.
10. Hayes, R.J., Petty, I.T.D., Coutts, R.H.A. and Buck, K.W. (1988). Nature 334, 179–182.
11. Elmer, J.S., Sunter, G., Gardiner, W.E., Brand, L., Browning, C.K., Bisaro, D.M. and Rogers, S.G. (1988). Plant Mol. Biol. 10, 225–234.
12. Brisson, N., Paszkowski, J., Penswick, J.R., Gronenborn, B., Potrykus, I. and Hohn, T. (1984). Nature 310, 511–514.
13. Hanley-Bowdoin, L., Elmer, J.S. and Rogers, S.G. (1988). Nucl. Acids Res. 16, 10511–10528.
14. Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987). EMBO J. 6, 3901–3907.
15. Bisaro, D.M., Hamilton, W.D.O., Coutts, R.H.A. and Buck, K.W. (1982). Nucl. Acids Res. 10, 4913–4922.
16. Bevan, M.W. (1984). Nucl. Acids Res. 12, 8711–8721.
17. Hamilton, W.D.O., Bisaro, D.M. and Buck, K.W. (1982). Nucl. Acids Res. 10, 4901–4912.
18. Lichtenstein, C.P. and Draper, J. (1985) in DNA Cloning : a Practical Approach, vol. 2 (ed. Glover, D.M.), 67–119 (IRL, Oxford).
19. Jefferson, R.A., Burgess, S.M. and Hirsh, D. (1986). Proc. Natl. Acad. Sci. USA 83, 8447–8451.
20. Petty, I.T.D., Coutts, R.H.A. and Buck, K.W. (1988). J. Gen. Virol. 69, 1359–1365.
21. Breathnach, R. and Chambon, P. (1981). Annu. Rev. Biochem. 50, 349–383.
22. An, G. (1986). Plant Physiol. 81, 86–91.
23. Jones, J., Dunsmuir, P. and Bedbrook, J. (1985). EMBO J. 4, 2411–2418.
24. Dean, C., Jones, J., Favreau, M., Dunsmuir, P. and Bedbrook, J. (1988). Nucl. Acids Res. 16, 9267–9283.
25. Deroles, S.C. and Gardner, R.C. (1988). Plant Mol. Biol. 11, 355–364.
26. Ward, A., Etessami, P. and Stanley, J. (1988). EMBO J. 7, 1583–1587.
27. MacDowell, S.W., Coutts, R.H.A. and Buck, K.W. (1986). Nucl. Acids Res. 14, 7967–7984.
28. Stanley, J. and Townsend, R. (1986). Nucl. Acids Res. 14, 5981–5998.
29. Lazarowitz, S.G. (1987). Plant Mol. Biol. Rep. 4, 177–192.

This article, submitted on disc, has been automatically converted into this typeset format by the publisher.