T-DNA promoter tagging in *Nicotiana tabacum*

A thesis presented in fulfilment of the requirements for the degree of Master of Philosophy in Genetics at Massey University, Palmerston North, New Zealand

Donald Haldane Kerr

1997

Abstract

Plant development is primarily controlled at the level of gene expression. In order to analyse this regulation it is necessary to isolate genes which are involved in organ development through cellular and tissue determination or which respond to environmental signals. Promoter tagging was chosen in order to identify genes potentially associated with plant development by their spatial and temporal pattern of expression. The introduction of a promoterless reporter gene tag allows the expression patterns of plant genes to be readily characterised.

A new series of promoter tagging vectors were constructed from the plasmid pPCV604 (Koncz, 1989). The selectable kanamycin resistance marker gene from pBin6 (Bevan, 1984) was cloned into pPCV604 to create pGT. The hygromycin phosphotransferase gene in pGT was then replaced with a promoterless β-glucuronidase (*gus*) gene coupled with octopine synthase termination sequence subcloned from pKiwi101a (Janssen and Gardner, 1989) creating pGTG. This binary transformation vector required the helper pRK replication functions of *Agrobacterium tumefaciens* strain GV3101. In order to bypass this restriction, the vector sequence of pBin19 was combined with the T-DNA of pGTG to create pBin19-GTG. The latter plasmid was found to have a higher *Agrobacterium tumefaciens*-mediated *Nicotiana tabacum* transformation efficiency in strain LBA4404 than pGTG in strain GV3101.

In both the pGTG and pBin19-GTG promoter tagging vectors the promoterless *gus* gene has an initiation codon 62 base pairs inside the T-DNA. This sequence includes translation termination codons in all three reading frames. Therefore, insertion of the T-DNA into a plant gene could lead to activation of the *gus* gene, under the control of the plant gene promoter, via transcriptional fusion.

Nicotiana tabacum leaf segments were transformed with pGTG or pBin19-GTG and transgenic plants selected on kanamycin. A population of 87 transgenic tobacco plants were fluorometrically screened for GUS activity in leaf and root material; 37% were found to contain GUS activity, indicating a high frequency of promoter tagging.

Two transgenic plants with root specific *gus* expression were analysed histochemically.

Progeny after self-fertilisation lacked GUS activity, though this was restored in progeny of one plant with 5-azacytidine treatment, suggesting involvement of methylation in the gene silencing. Southern hybridisation, inverse PCR cloning of T-DNA flanking sequences and segregation on kanamycin indicated the presence of multiple T-DNA copies within the primary transformants. Furthermore, inverse PCR sequence from one plant indicated multiple and truncated T-DNA insertions at one or more loci.

A further population of transformed plants was generated with pBin19-GTG and histochemically screened for GUS activity in roots (14 positive from 147 tested), shoots (27 positive from 147) and floral organs (14 positive from 56). Overall, combining results from all plant organs tested, an average of 33% of plants were found with GUS activity in one or more organs. A diverse range of patterns of *gus* expression were observed and described including patterns involving root branching.

Forty four plants from this population were analysed for T-DNA copy number via Southern hybridisation with a *gus* probe (right border junction T-DNA) and *npt*II probe (central T-DNA). Multiple copies were frequently found with an average of 3.3 T-DNA copies per transgenic plant. Overall, an average of 11% of T-DNA insertions were found to be involved in *gus* activation.

Comparison of the fluorometric (37% positive, 87 plants tested) and histochemical (22% positive, 147 plants tested) screens for GUS activity in root and shoot material was discussed and it is suggested that further care is needed in assigning promoter tagging hits from fluorometric screening.

Variable expression was observed with promoter tagged genes. It is suggested that further research is required to determine whether this variation was due to silencing, perhaps by methylation, or was a result of the tagged promoters' normal expression patterns.

Acknowledgements

I would like to thank my supervisors, Dr Derek White, Professor Barry Scott and latterly Professor Paula Jameson, for their patience and advice during this research study.

I am grateful to Dr Csaba Koncz for the plasmid pPCV604 which made this study possible and also for sharing unpublished results.

I acknowledge with thanks the receipt of essential personal funding from a Massey University Scholarship, a University Grants Committee Scholarship, the William Georgetti Trust, Grasslands DSIR and finally an Assistant Lectureship within the Plant Biology and Biotechnology Department.

I have been privileged to study within the Plant Molecular Genetics Laboratory at AgResearch Palmerston North (formerly Grasslands DSIR) and would like to thank the scientists, technicians and fellow students there for their support, both scientific and personal.

I am also grateful to the staff and students within Scott Base (Molecular Genetics Unit) and in the Plant Biology and Biotechnology Department for occasional assistance, advice and use of equipment, especially Lynnaire Oswald for assistance with tissue culture and glasshouse work in the last eighteen months.

Thanks also to my friends and family who have supported me, especially my parents Joan and Barry Kerr and friends from the congregation of All Saints' Church, Palmerston North.

I wish to credit Peter Spring, formerly of DSIR Science Centre Palmerston North, for the photographs in Figures 16, 17, 18 **A**, **B** and **G**, 19, 20 **A** and 22 and Lynnaire Oswald, Plant Biology and Biotechnology Department, for the majority of the photographs in Figure 31.

Table of Contents

List of Figures

List of Tables

Abbreviations

AB	*Agrobacterium* minimal media, Section 2.1.1
Ap	ampicillin
CTAB	hexadecyltrimethylammonium bromide
EDTA	ethylene diamine tetra acetic acid, sodium salt unless specified
Gm	gentamycin
gus	β-glucuronidase gene
GUS	β-glucuronidase enzyme
hpt	hygromycin B phosphotransferase
kb	kilobase pairs
Km	kanamycin
LB	left border of T-DNA, a 25 base pair conserved sequence
LB	Luria Bertani bacterial media, Section 2.1.1
MS	Murashige and Skoog plant media, Section 2.4.2
MU	methyl umbelliferone
MUG	methyl umbelliferone glucuronide
Nic	*Nicotiana* media, Section 2.4.2
nos	nopaline synthase gene from *Agrobacterium tumefaciens* T-DNA
npt	neomycin phosphotransferase gene, conferring kanamycin resistance
ocs	octopine synthase gene from *Agrobacterium tumefaciens* T-DNA
P_{35S}	promoter region of the 35S cauliflower mosaic virus transcription unit
PCR	polymerase chain reaction
P*mas*	promoter region of *mas* gene
P*nos*	promoter region of *nos* gene
RB	right border of T-DNA, a 25 base pair conserved sequence
Rf	rifampicin
Ri	root inducing, referring to *Agrobacterium rhizogenes* plasmid
rpm	revolutions per minute
SOB	bacterial media, Section 2.1.1
SOC	bacterial media, Section 2.1.1
SDS	sodium dodecyl sulphate
SSC	saline sodium citrate (20xSSC: 3 M NaCl, 0.3 M trisodium citrate)

SUDS	an agarose gel loading buffer also containing detergent suitable for halting restriction endonuclease enzyme activity, Section 2.3.12
Tc	tetracycline
T-DNA	transferred DNA, essentially the section of DNA bounded by left and right border sequences that is inserted into the plant genome by *Agrobacterium*
TE	buffer, 10 mM Tris-HCl, 1 mM EDTA, pH 8.0
Ti	tumour inducing, referring to *Agrobacterium tumefaciens* plasmid
Tp	trimethoprim
TY	Tryptone and yeast extract bacterial media, Section 2.1.1
vir	virulence genes on *Agrobacterium* Ti plasmid
XGal	5-bromo-4-chloro-3-indolyl β-D-galactoside
XGluc	5-bromo-4-chloro-3-indolyl β-D-glucuronide (cyclohexylammonium salt)

Other standard abbreviations are as used by Biochemical Journal (1992)

1.0 Introduction

1.1 Gene regulation

Plant gene expression is the driving force behind plant growth and development. The combination of the choice of genes expressed and the spatial and temporal control of that expression create the characteristic living organisms that are recognised as plants.

Advances in understanding the control of gene expression in plants has relied on two fundamental strategies, the cloning of plant genes, and the production of transgenic plants. Together these have allowed the analysis of gene structure at the nucleotide sequence level and the testing of modified regulatory elements within transgenic plants to determine their role in gene expression.

A typical plant nuclear gene (Waugh and Brown, 1990; Kuhlemeier, 1992; Robinson *et al.*, 1993) coding for a protein consists of upstream regulatory elements, then a TATA box which directs RNA polymerase II to begin transcription approximately 30 base pairs downstream. At the 3′ end of the gene are transcription termination sequences which to date remain poorly defined.

The transcribed sequence is capped with a 7-methyl guanosine residue through a 5′-5′ linkage to the triphosphate group at the end of the RNA molecule. Near the 3′ end of the RNA are one or more signals, usually AAUAAA, which promote polyadenylation about 20 bases further downstream. Not all RNA molecules are polyadenylated; one major class of mRNA within this category are the histone genes. Within the transcribed sequence may be further signals for the removal of introns. The intron sequence begins with a GT dinucleotide and finishes with AG. Another important splicing signal is a high AU content within the intron.

The coding region of the transcript begins with the translation initiation codon AUG near the 5′ end of the mRNA and, after an open reading frame, terminates with a stop codon (UAG, UAA or UGA).

While gene expression could be regulated at any point from transcription, RNA

processing and turnover to translation, as well as control of gene activity via the rate of protein processing and turnover, it is the rate of transcription which is the primary regulatory event. Hence the majority of studies of plant gene regulation have focused on the control elements upstream of the TATA box.

About 75 base pairs upstream of the transcription initiation site is often a CAAT or AGGA box. These sequences are thought to assist RNA Polymerase II binding at the TATA box.

Further upstream, within about 1 kilobase, are various sequences that act as enhancers or silencers via protein binding and provide tissue and/or stimulus specificity to gene expression. The complex interactions of these sequences with transcription factors and regulatory proteins are thought to determine the final pattern of gene expression. Following gene cloning, the function of these sequences are analysed using transgenic plants and reporter genes.

1.2 Gene cloning

There are two main approaches to gene cloning, those relying on standard molecular techniques to probe cDNA libraries (and subsequently genomic libraries) and genetic methods utilising mutants (Walden and Schell, 1990). More recently genes have been identified following direct sequencing of large portions of nuclear genomes or expressed sequences (Gardiner and Mural, 1995).

The standard molecular techniques include (i) cross-hybridisation with a homologous gene from another organism, (ii) antibody detection of a protein synthesised by an expression library, (iii) hybridisation with a degenerate oligonucleotide probe determined from protein sequence or PCR with two such oligonucleotides, (iv) differential screening using labelling of mRNA from two different cell types to seek clones representative of just one of the cell types and (v) subtractive generation of a library removing sequences homologous to mRNA from other cell types thereby enriching the library for expressed genes representative of one cell type (Meyerowitz, 1992; Robinson *et al.*, 1993).

These techniques all require some knowledge of the DNA or protein sequence, or of the pattern of gene expression, and therefore are unsuitable for cloning genes revealed by mutant phenotypes. When a mutation has been generated by chemical means or by irradiation then a map or positional cloning approach must be used to clone the gene. For example, a restriction fragment polymorphism map can be used to initially locate the gene and provide a reference point from which to move along overlapping clones (yeast artificial chromosomes) to the gene locus (Meyerowitz, 1992).

Another approach for cloning a mutant gene without mapping or knowledge of the gene products is genomic subtraction (Meyerowitz, 1992). This involves hybridising wild type DNA with the mutant DNA and amplifying by PCR unhybridised fragments.

Less rewarding so far is 'shotgun' transformation of a plant genomic library into a suitable host (often mutant) and screening for a new phenotype (or complementation). This has been successful with particular reporter constructs selecting for an entire plant promoter (Herman *et al.*, 1986) and for enhancer elements (Ott and Chua, 1990) but not for unknown entire plant genes, though trials have been successful with marker genes (Walden and Schell, 1990).

Another approach is to generate mutations by randomly inserting a known fragment of DNA into the genome. When the marker DNA inserts into an active gene, the gene will usually be disrupted resulting in a mutant phenotype. The marker DNA facilitates the cloning of the surrounding sequence, and hence the gene of interest. This technique is known as insertional mutagenesis or gene tagging.

1.3 Gene tagging

Gene tagging has been widely used to assist in cloning genes in organisms ranging from bacteria (Kleckner, 1981) to the fruit fly *Drosophila* (Bingham, Levis and Rubin, 1981) and mammals (Gridley *et al.*, 1987) as well as in plants (Walbot, 1992). The tagging DNA can then be used as a marker to allow the isolation of the flanking DNA sequence, which in turn can be used to clone the wild type gene.

The nature of the DNA used for gene tagging has varied among different biological systems. However it must be an element that has the ability to recombine into the genome, preferably at random. In bacteria, transposons such as Tn5 (Kleckner, 1981) are readily mobile into both genomic and plasmid DNA and carry a selectable antibiotic resistance gene. They are usually introduced on a plasmid incapable of maintenance in that strain. Hence selection for the antibiotic resistance carried by the transposon ensures a transposition event has occurred.

In *Drosophila*, tagging has been conducted with naturally occurring mobile elements such as *copia* and P-element (Cooley *et al.*, 1988). The first successful cloning of a gene was that of the *white* locus (Bingham, Levis and Rubin, 1981). They utilised a pre-existing strain with the mobile element *copia* known to have created a mutant allele.

Mice lack a suitable cloned transposable element, so gene tagging has been conducted by direct transformation with foreign DNA or by retroviral infection (reviewed by Gridley *et al.*, 1987).

In plants the analogous tagging elements are the endogenous transposons and also the transferred DNA (T-DNA) introduced into the plant genome by *Agrobacterium*.

1.4 Transposon tagging

The transposons in *Zea mays* and *Antirrhinum majus* have been well characterised at the molecular level (Gierl and Saedler, 1992) and used for the cloning of numerous genes in these two organisms (reviewed in Walbot, 1992).

Two general strategies have been used to clone genes using transposons for insertional mutagenesis (Gierl and Saedler, 1992; Walbot, 1992). Targeted transposon tagging seeks to clone a particular gene whose map location is previously known. As transposable elements appear to preferentially insert into nearby chromosomal locations (Walbot, 1992) this requires the insertion of an active transposon near the locus of interest. Non-targeted transposon tagging relies on random insertions being

able to mutate a gene or genes of unknown position. This requires active transposable elements throughout the genome in order to be able to hit most or all chromosomal sites.

While transposon tagging has been very successful in both *Z. mays* and *A. majus*, it has taken considerable time for this strategy to be applied in other plant species (Shields, 1993). As transposable elements have not been well characterised in other plant species, transposon tagging was attempted with a cloned element introduced from another species. This technique is known as heterologous host transposon tagging (Hehl, 1994). Baker *et al.* (1986) confirmed the validity of this approach by demonstrating that the *Activator* (*Ac*) element from *Z. mays* could transpose after it was introduced into tobacco via *Agrobacterium*-mediated transformation. However, it was not until 1993 that the first report of successful heterologous transposon tagging appeared. Chuck *et al.* (1993) cloned the *Ph6* gene in *Petunia* after insertion of an introduced *Ac* element. The *Ph6* gene influences floral colour via an effect on the pH of the vacuole. Hence revertants caused by re-excision of the *Ac* element during somatic development of the flower caused a conspicuous variegation in petal colour. This demonstrates one of the advantages of transposon tagging over T-DNA gene tagging, that the cause of the mutational event can be identified as a transposon by the distinctive reversion effects.

Walbot (1992) and Coupland (1992) discuss the advantages and disadvantages of heterologous transposon tagging versus T-DNA gene tagging, especially with regard to *Arabidopsis*. Difficulties have occurred with transposon tagging, especially due to a low frequency of transposition, and hence a low insertional mutagenic frequency, though modifying the promoter and untranslated leader sequence of the *Ac* element transposase greatly increases this rate. Hence a large number of potential mutagenic insertions can be generated in a self cross. This gives a significant advantage over T-DNA tagging where each individual tagging event must arise from an individual transformed plant. Another advantage of transposon tagging is that the *Agrobacterium*-mediated transformation process leads to a significant number of non-linked mutations which complicate the analysis as they cannot be isolated using the T-DNA probe (Koncz *et al.*, 1992).

1.5 T-DNA gene tagging

The first reports of T-DNA gene tagging (Teeri *et al.*, 1986; Andre *et al.*, 1986) used promoterless antibiotic resistance genes for positive selection of tagging events. Thus tagging events were only seen when the T-DNA insertion into an plant gene was such that the antibiotic resistance gene was expressed under the control of a plant gene promoter. This specialisation of gene tagging is now known as promoter tagging (or promoter trapping) and is reviewed below (Section 1.6).

The most significant study of T-DNA gene tagging has been the work of Feldmann and colleagues (Feldmann and Marks, 1987; Feldmann *et al.*, 1989; Feldmann *et al.*, 1990; Feldmann, 1991; Feldmann, 1992; Forsthoefel *et al.*, 1992). They have developed a system for *Agrobacterium*-mediated transformation of *Arabidopsis* seed, using a T-DNA carrying a kanamycin resistance marker. Among 13000 independent kanamycin resistant transformed lines screened for a wide range of mutations, approximately 20% have mutant phenotypes (Forsthoefel *et al.*, 1992).

Among the first genes cloned following mutant characterisation have been glabrous (Herman and Marks, 1989) and agamous (Yanofsky *et al.*, 1990). As the seed transformation lines have been made freely available to the whole scientific community the number of different mutant screens and consequently the number of mutations found continues to expand.

One major advantage of this seed transformation system is that the T-DNA is introduced into plants without a requirement for laborious tissue culture. As well, it was thought that this would also avoid mutational events (somaclonal variation) occurring during regeneration from tissue culture (Feldmann and Marks, 1987). However, subsequent analysis of mutations found within transformed lines of *Arabidopsis* has revealed that approximately 50% are not linked to T-DNA insertions (Forsthoefel *et al.*, 1992).

Koncz *et al.* (1992) compared the mutational frequency from the seed transformation experiments (8000 lines) to that of tissue culture derived transformed plants (1340 plants) and found the frequencies to be similar, and comparable to chemical

mutagenesis. Also, they estimate that the mutations found in both transformation systems are only linked to T-DNA insertions in 10 to 40% of the plants.

One potential explanation (Koncz *et al.*, 1992) is that the process of T-DNA insertion may lead to deletions or chromosome breakages at sites where the T-DNA integration fails and a DNA repair system is activated, hence giving rise to a mutations unlinked to a T-DNA. This prediction has not been proved, but evidence is accumulating. Marton *et al.* (1994) found a greatly increased mutation frequency in regenerating *Nicotiana* protoplasts when *Agrobacterium*-mediated transformation occurred. Studies in *Petunia* have shown that the initial concentration of transferred-DNA, detected by Southern hybridisation, is high, but rapidly drops (Virts and Gelvin, 1985) and that the T-DNA is initially transcribed in a large number of cells, though only a small fraction of these cells appear to become stably transformed (Janssen and Gardner, 1989).

The frequent occurrence of mutations unlinked to T-DNA integration sites is problematic, but the large numbers of plants screened means that there are plentiful linked mutations to isolate and unlinked mutations can still be cloned, via restriction fragment length polymorphism mapping and chromosome walking (Koncz *et al.*, 1992). A variation of gene tagging called promoter tagging, utilising a reporter gene to provide immediate information on the actual tagged locus, avoids the problem of unlinked mutations.

1.6 Promoter tagging

A more sophisticated gene tagging approach has been to incorporate a promoterless reporter gene at one end of the mobile DNA element. Then insertion into an active gene will lead not only to a mutagenic event but also potentially to the activation of the reporter gene. This will occur when the insertion aligns the reporter gene coding sequence in the same orientation as the tagged gene, when the insertion has not disrupted the promoter region and when there are no transcriptional or translational fusion difficulties. This strategy of gene tagging was first applied in bacteria (Kleckner, 1981). Typical is the use of Tn3HoHo1 (Stachel *et al.*, 1985) which contains a promoterless β-galactosidase gene immediately inside the repeat sequence

which forms the boundary of the transposing element.

As mentioned above, promoter tagging (also known as promoter trapping) was first described in plants in 1986 with a promoterless neomycin phosphotransferase gene (*npt*II) immediately downstream of a T-DNA right border transformed into tobacco. Teeri *et al.* (1986) obtained eight kanamycin resistant plants after selection during regeneration, at a frequency of 0.0002% while Andre *et al.* (1986) reported 147 kanamycin resistant plants at a frequency of 5%. The lower frequency scored by Teeri *et al.* (1986) may have been due to their vector lacking a T-DNA left border.

Unfortunately, selection for a promoterless antibiotic resistance gene leads to a greatly increased T-DNA copy number within the plant genome (5 to 20 copies found in Southern hybridisation analysis of over 200 transformants) compared with that after normal selection for a constitutively expressed antibiotic resistance gene (average 1.5 copies, Koncz, 1989; Koncz *et al.*, 1992). This would hinder any attempt to clone flanking promoter sequences. On the other hand, the effectiveness of the technique was demonstrated, generating insertions into genes in the primary transgenic plant without the need to self-fertilise to produce recessive mutants which is necessary with transposon tagging and T-DNA gene tagging.

An alternative promoter tagging strategy which retains the utility of screening in the primary transgenic but avoids selection pressure is to use a non-selected promoterless reporter gene, such as β-glucuronidase (*gus*A; Kertbundit *et al.*, 1991; Fobert *et al.*, 1991; Christey *et al.* 1993; Lindsey *et al.*, 1993; Topping *et al.*, 1994), luciferase (*lux*A, with *lux*B constitutively expressed; Jiang *et al.*, 1992) or β-galactosidase (*lacZ*; Teeri *et al.*, 1989; Suntio and Teeri, 1994) within a T-DNA. The T-DNA also contains a normal antibiotic resistance selectable marker gene which allows selection for transformed cells. Following regeneration under standard selection a population of transgenic plants are screened for the promoterless *gus*, *lux* or *lacZ* expression.

A result common to many of the studies (summarised in Koncz *et al.*, 1992; also Lindsey *et al.*, 1993) is that the frequency of reporter gene expression is high, regardless of the species' ratio of single copy to repetitive genomic DNA. Hence it seems likely that the T-DNA is integrating preferentially into transcribed DNA

sequences, which makes gene tagging a profitable technique for isolating new genes not only in *Arabidopsis* with its low repeat content, but also in other plants with much greater proportions of repetitive DNA.

Koncz *et al.* (1989) demonstrated that both transcriptional and translational fusions could lead to expression of the promoterless reporter gene and also that dicistronic units functioned effectively, at least as long as the initiation codon of the second gene was within close proximity to the first gene's stop codon. Hence a promoterless reporter gene could function in a dicistronic unit, rather than through translational fusion.

1.7 Enhancements of gene tagging

An assortment of gene tagging vectors with specific functional targets have been constructed and used in a variety of eukaryotic species (Gossler and Zachgo, 1993). These include enhancer tagging (or trapping) in which the reporter gene retains a minimal promoter with (only) a TATA box. Insertion near enhancer elements will lead to expression of the reporter. This has been successfully tested in potato by Goldsbrough and Bevan (1991) with a T-DNA containing a *gus* gene with a minimal 35S Cauliflower Mosaic Virus promoter. It is likely that the pattern of gene expression revealed by the reporter will reflect the tissue specificity of the activating enhancer elements.

Exon trapping (also known as gene trapping) uses a reporter gene without promoter or initiation codon (Gossler and Zachgo, 1993). Instead, the reporter is positioned immediately downstream of an intron splice acceptor site. Hence the reporter gene will only be expressed on insertion into an intron sequence. A recent example of this technique was the report (Springer *et al.*, 1995) of the cloning of the gene *Prolifera* using an *Ac/Ds* (Activator/Dissociation) maize transposon system in *Arabidopsis*. The *gus* reporter gene revealed that the tagged locus, while leading to megagametophyte death due to insertional mutagenesis, was also expressed in dividing cells throughout the plant.

Another refinement of gene tagging is activation tagging (reviewed by Walden *et al.*, 1994a) which uses a T-DNA containing multiple enhancer sequences. These may increase the transcription rate of genes near the T-DNA insertion site, giving rise to dominant mutations. By specific selection under conditions normally not conducive to growth, such as regeneration on auxin deficient media (Hayashi *et al.*, 1992) or regeneration on cytokinin deficient media (Walden *et al.*, 1994a) or with an inhibitor of polyamine biosynthesis (Fritze *et al.*, 1995), mutants affecting metabolism or function in each of the respective hormone classes have been isolated. A protoplast transformation system has been used to generate sufficient numbers of transformants. With a typical transformation efficiency of 20%, $30x10^6$ protoplasts generated 12 auxin independent plants (Hayashi *et al.*, 1992) and eight mutants in polyamine metabolism (Walden *et al.*, 1994a).

The activation tagging with auxin independent regeneration has lead to the characterisation of *axi 1*, expression of which is normally activated by high auxin levels and leads to cell division (Walden *et al.*, 1994b) The *axi 1* transcript initiates 6 kb upstream of the activating T-DNA insertion (Walden *et al.*, 1994a). This demonstrates the effectiveness of this procedure which therefore appears likely to have a much larger target sequence within the plant genome for a potential tagging event than either promoter tagging or insertional mutagenesis.

Another improvement on T-DNA tagging likely to lead to developmentally useful mutants is suicide tagging (Koncz *et al.*, 1992). This variation of promoter tagging uses a promoterless gene which, when expressed, will lead to cell death, such as the barnase gene from *Bacillus amylolyquefaciens*. This could lead to transgenic plants with 'missing feature' mutants (Koncz *et al.*, 1992).

Whichever variation on T-DNA tagging is studied, all require an initial *Agrobacterium*-mediated transformation step.

1.8 *Agrobacterium*-mediated transformation

Agrobacterium-mediated transformation utilises a unique bacterial plant interaction in which DNA is transferred from the bacteria and stably inserted into the plant nuclear genome (reviewed by Ream, 1989). *A. tumefaciens* and *A. rhizogenes* transformation result in tumours and hairy root growths respectively in susceptible, mainly dicotyledonous, host plants. This ability resides on the large, approximately 150 to 250 kilo base plasmids named Ti (tumour inducing) and Ri (root inducing). A part of that plasmid, the T-DNA (transferred DNA) of about 20 kb is incorporated into the plant genome and carries genes involved in hormone metabolism and opine synthesis. The T-DNA is bordered at each end by a 25 base pair imprecise direct repeat sequence. The right border is essential for transfer of the T-DNA and can function without a corresponding left border (Wang *et al.*, 1984). The plant/T-DNA boundary is usually found at the inside edge of the right border sequence, but is much less precise at the left edge of the T-DNA (Bakkeren *et al.*, 1989). In some strains the T-DNA is separated into two distinct regions on the Ti or Ri plasmid. These are called T_L and T_R (transferred-left and transferred-right) and both are bracketed by the 25 base pair repeats (Ream, 1989).

Studies of T-DNA integration sites in the plant genome (Matsumoto *et al.*, 1990; Gheysen *et al.*, 1991; Mayerhofer *et al.*, 1991) have revealed the absence of all but very short sequence homologies. It is vital that the T-DNA integration site should lack any particular sequence preference for the success of T-DNA gene tagging as all genes within the genome should be available for T-DNA insertion.

The molecular mechanism of T-DNA transfer and recombination into the plant DNA (reviewed by Kado, 1991 and Hooykaas and Schilperoort, 1992) has yet to be fully resolved. Specific plant phenolic compounds, produced by wounded plant cells, induce the virulence (*vir*) genes on the Ti or Ri plasmid via the activity of the two component regulatory system made up of the *virA* and *virG* gene products, which are constitutively expressed.

VirA is found in the bacterial inner membrane and binds phenolic compounds such as acetosyringone, after which (deduced by comparison to other two component

regulatory systems) it activates VirG by phosphorylation, which in turn binds to *vir* promoters and activates their transcription.

VirD1 and VirD2 together have an endonuclease activity that specifically nicks DNA within the 25 base pair border repeat element. Both single stranded (bottom strand, named T-strand) and, at a lower frequency, double stranded T-DNA molecules can then be detected, though it is thought that the T-strand is the molecule transferred to the plant. VirD2 remains attached to the T-strand at its 5′ end which protects the DNA from exonucleases and may aid movement into the nucleus after transfer as it contains nuclear targeting signals. The VirE2 gene products also associate with the protein acting as single stranded DNA binding proteins which coat the T-strand.

The initial steps of nicking and T-strand synthesis and many of the proteins involved appear homologous to those of the rolling circle replicative transfer of plasmids occurring in bacterial conjugation.

The *virC1* product binds to a sequence, named overdrive (Peralta *et al.*, 1986) found external to the right border which acts as an enhancer for both T-strand synthesis and transformation efficiency.

The products of the *virB* operon are also essential for transfer. Many of these proteins are associated with the bacterial membrane and probably form a conjugative pilus. Other *vir* genes (*virF* and *virH*) and further genes, both of Ti and chromosomal origin, have effects on host specificity.

This system of *Agrobacterium*-mediated transformation has been manipulated for plant genetic engineering by making use of the discoveries that the only DNA sequence required for recognition of the T-DNA for its transfer are the 25 base pair border repeats (Wang *et al.*, 1984) and that these can function in *trans* with respect to the *vir* genes (Hoekema *et al.*, 1983).

Hence transformation vectors can be constructed with only the right and left borders, without the opine and hormone metabolism genes, on smaller easily manipulated plasmids able to replicate in both *E. coli* and *Agrobacterium*. Combining these

plasmids with a Ti plasmid carrying the *vir* genes and usually disarmed by deletion of the T-DNA creates a binary transformation system (for example see Bevan, 1984). The transformation vector is then suitable for introducing any gene or genes, between the right and left borders, for study within plants. This system has been the basis of the T-DNA promoter tagging conducted in this study.

1.9 Aims

The aim of this study was to construct a T-DNA promoter tagging vector (Section 1.10) with a promoterless *gus* reporter gene and to use this vector to analyse promoter tagging in *Nicotiana tabacum*. A population of transgenic plants with independent T-DNA insertions would be generated. Screening of these plants for GUS activity and for the number of T-DNA insertions would provide information on the frequency of promoter tagging and the range of patterns of gene expression revealed by this technique.

It was anticipated that it would be possible to clone plant genomic DNA flanking one or more of the T-DNA insertions using one of three techniques for rescuing T-DNA flanking sequence, namely, inverse PCR (Triglia *et al.*, 1988; Ochman *et al.*, 1988), single sided ligation mediated PCR (Fors *et al.*, 1990) or plasmid rescue. Plasmid rescue would require specific elements within the T-DNA (refer Section 1.10). Subsequently, suitable cloned fragments could be investigated as to their promoter activity. The final demonstration of this would be by transforming a fragment into tobacco within a reporter gene construct and recreating the pattern of *gus* expression seen in the original promoter tagged plant.

1.10 Promoter tagging vector

The most important features required in the plant T-DNA promoter tagging vector were an easily assayed promoterless reporter gene with its start codon just downstream of the right border element, a separate selectable marker gene to generate and distinguish transformed plants and a system to allow the rescue of T-DNA and

flanking plant genomic sequences. These features all needed to be accommodated within the T-DNA of a binary vector suitable for *Agrobacterium*-mediated plant transformation.

The promoter tagging vector pPCV604 (Figure 1; Koncz, 1989) was available courtesy of Dr Csaba Koncz (Max Planck Institute) as the starting point for the construction of a new gene tagging vector. The plasmid pPCV604 is a derivative of the pPCV plasmid series which has been described by Koncz and Schell (1986). It is a binary T-DNA vector which contains a promoterless hygromycin B phosphotransferase (*hpt*) gene immediately downstream of the T-DNA right border. Hence it can be used to analyse gene tagging via selection for transformed plants on hygromycin.

The plasmid pPCV604 was constructed by Koncz (1989) by inserting a *Bam*HI restriction fragment containing the *hpt* protein coding sequence into unique *Bam*HI-*Bcl*I sites in pPCV002, and then adding downstream sequences from the nopaline synthase (*nos*) gene into unique *Hin*dIII-*Bam*HI sites to provide 3′ polyadenylation signals for the *hpt* gene.

The *hpt* translation initiation codon is 62 base pairs downstream of the 25 base pair right border sequence which defines the edge of the T-DNA. The intervening sequence contains stop codons in all three reading frames. This prevents translational read through so that the reporter gene activity will not be influenced by any N-terminal peptide fusion as might be expected in a gene tagging event internal to the coding sequence of a plant gene.

There is no TATA box sequence between the right border and the reporter gene start codon so that transcription of *hpt* must be the result of a transcriptional fusion downstream of a plant gene transcription initiation site. Hence this vector is a promoter tagging vector rather than enhancer tagging (requiring own TATA box).

The plasmid pPCV604 also contains other features useful in a gene tagging vector including a beta-lactamase gene and an origin of replication from the plasmid pBR322. These are contained inside the T-DNA so that they will be transferred into the plant genome with the promoter tagging reporter gene, and will enable the cloning of the

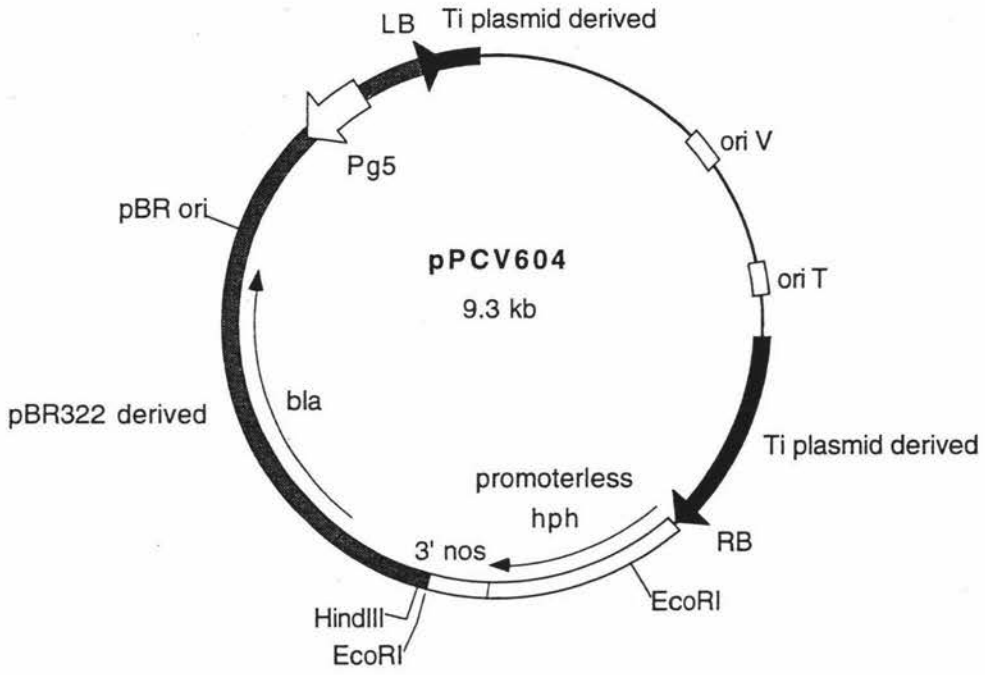


Figure 1. Physical map of pPCV604.

Information from Koncz (1989).

Abbreviations: LB and RB, left and right borders of T-DNA; pBR ori, replication origin from pBR322; *hpt*, coding region of hygromycin phosphotransferase gene; *bla*, β-lactamase gene; Pg5, promoter T-DNA gene 5.

T-DNA and flanking sequence using plasmid rescue.

However, this promoter tagging vector does not separate the functions of the reporter gene and the selectable marker. This was regarded as being a negative feature for three reasons. First, it is likely that only T-DNA insertions resulting in moderate to strong levels of *hpt* expression in callus and shoot material would survive initial selection on hygromycin. Hence, many successful T-DNA insertions into plant genes not expressing within these tissues and at the time of selection could be lost. Second, it has been reported (Koncz, 1989; Koncz *et al.*, 1992) that positive selection for promoter tagging results in a higher frequency of transformants with multiple copies of active gene insertions. This would greatly complicate the analysis and cloning of flanking plant DNA sequences involved in insertion events. Third, it would not be practical to monitor *hpt* gene expression in order to ascertain fine details of expression patterns resulting from promoter tagging events.

Three modifications to pPCV604 were envisaged to improve its usefulness as a promoter tagging vector. An independent constitutively expressed selectable marker gene (*npt*II) could be added so that all transgenic plants could be selected on the antibiotic kanamycin without any (known) selection pressure on the putatively independent gene tagging events. Further, the promoterless hygromycin resistance gene used for gene tagging in Koncz's plasmid could be replaced with a promoterless β-glucuronidase (*gus*) gene. This critical step would allow gene tagging events to be observed much more easily, through analysing the activity of the *gus* gene. GUS activity can be monitored both histochemically, yielding cell specific expression information, and fluorometrically, to quantify the expression. Finally, the strain dependence of the plasmid needed to be modified so that it could be used in a much wider range of *Agrobacterium* strains with practical benefits for plant transformation experiments.

2.0 Methods and materials

2.1 Bacterial work

2.1.1 Bacterial media

LB 1% (w/v) tryptone, 1% (w/v) NaCl, 0.5% (w/v) yeast extract, adjusted to pH 7.5 with 2 M NaOH (Maniatis *et al.*, 1982).

TY 0.5% (w/v) tryptone, 0.3% (w/v) yeast extract, 0.087% (w/v) $CaCl_2$ (Beringer, 1974).

AB 0.3% (w/v) K_2HPO_4, 0.1% (w/v) NaH_2PO_4, 0.1% (w/v) NH_4Cl, 0.03% (w/v) $MgSO_4.7H_2O$, 0.015% (w/v) KCl, 0.001% (w/v) $CaCl_2$, 0.00025% (w/v) $FeSO_4.7H_2O$, 0.5% (w/v) glucose (Chilton *et al.*, 1974)

SOB 2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$ (filter sterilised), 10 mM $MgSO_4$ (filter sterilised) (Hanahan, 1983).

SOC SOB media with 20 mM glucose (filter sterilised) (Hanahan, 1983).

2.1.2 Bacterial strains and DNA vectors

Bacterial strains and DNA vectors used in this study are listed in Table 1.

2.1.3 Bacterial strain maintenance

Bacterial strains were stored long term in deep frozen (-70°C) 15% (v/v) glycerol stocks, prepared by growing overnight broths with antibiotic selection, as appropriate, and mixing 700 µl of culture with 300 µl of 50% (v/v) glycerol. When in constant use, strains were maintained by restreaking on solid media with the appropriate antibiotic added. *E.coli* were grown at 37°C using LB media either as broth or solid (1.5% w/v agar) media, and *Agrobacterium* were grown at 26°C on solid AB or broth TY media. Plates were stored short term at 4°C.

Table 1. Bacterial strains and DNA vectors.

Strain or replicon	Relevant characteristics	Source or reference
Bacterial strains		
Escherichia coli		
HB101	F⁻Δ(*gpt-pro*A)62 *leu sup*E44 *ara*14 *gal*K2 *lac*Y1 Δ(*mcr*C-*mrr*) *rps*L20 (Sm^R) *xyl*-5 *mtl*-1 *rec*A13	Boyer and Roullard-Dussoix, 1986
DH5α	F⁻ *end*A1 *hsd*R17 (r_k^-, m_k^+) *sup*E44 *thi*-1 λ⁻ *rec*A1 *gyr*A96 *rel*A1 *deo*R Δ(*lac*ZYA-*arg*F)U169 ϕ80d*lacZ*ΔM15	Hanahan, 1983
DH5αMCR	F⁻ *end*A1 *sup*E44 *thi*-1 λ⁻ *rec*A1 *gyr*A96 *rel*A1 *deo*R Δ(*lac*ZYA-*arg*F)-U169 ϕ80d*lac*ZΔM15 *mcr*A Δ(*mrr hsd*RMS *mcr*BC)	Grant *et al.*, 1990
S17-1	Tp^R Sm^R *pro*, chromosomally integrated plasmid RP4 2-Tc::Mu-Km::Tn7, *hsd*R⁻, *hsd*M⁺	Simon *et al.*, 1983
DH10b	F⁻ *ara*D139 Δ(*ara, leu*)7697 Δ*lac*X74 *gal*U *gal*K *rps*L *deo*R ϕ80d*lac*ZΔM15 *end*A1 *nup*G *rec*A1 *mcr*A Δ(*mrr hsd*RMS *mcr*BC)	Grant *et al.*, 1990; Lorow and Jessee, 1990
Agrobacterium		
LBA4404	*Agrobacterium tumefaciens* octopine strain carrying a large cryptic plasmid and pAL4404	Ooms *et al.*, 1982; Hoekema *et al.*, 1983
GV3101	*Agrobacterium tumefaciens* C58C1 Rf^R carrying pMP90RK	Van Larebeke *et al.*, 1974; Koncz and Schell, 1986
DNA vectors		
***Agrobacterium* replicons**		
pAL4404	disarmed derivative of octopine Ti plasmid pTiAch5 through deletion of the T-DNA, binary *Agrobacterium*-meditated plant transformation vector providing *vir* genes	Ooms *et al.*, 1982; Hoekema *et al.*, 1983
Dual *Agrobacterium* and *E.coli* replicons		
pMP90RK	Gm^R, Km^R derivative of pTiC58 with T-DNA deleted and insertion of pRK2013 sequence with plasmid maintenance and mobilisation functions, binary *Agrobacterium*-meditated plant transformation vector providing *vir* genes	Koncz and Schell, 1986
pPCV604	Ap^R promoter tagging binary T-DNA vector with promoterless *hpt* gene,	Section 1.10; Koncz, 1989
pBin19	Km^R binary plant transformation vector with T-DNA	Bevan, 1984
pBin6	Km^R binary plant transformation vector with T-DNA	Bevan, 1984
pRK2013	Km^R helper plasmid for triparental matings	Figurski and Helinski, 1979; Ditta *et al.*, 1980
pGT	Ap^R	this study, Section 3.1.1
pGTG	Ap^R	this study, Section 3.1.2
pBin19-GTG	Ap^R	this study, Section 3.1.3

Table 1 continued.

Strain or replicon	Relevant characteristics	Source or reference
DNA vectors		
***E.coli* replicons**		
λ	used for molecular weight markers and DNA concentration standards	purchased from Promega Corporation
pUC8	Ap^R	Vieira and Messing, 1982
pBR322	Ap^R Tc^R	Bolivar *et al.*, 1977
pUC18	Ap^R	Yanisch-Perron *et al.*, 1985
pUC19	Ap^R	Yanisch-Perron *et al.*, 1985
pKIWI101a	Ap^R containing P*mas*-P_{35S}(truncated)-*gus*-3′*ocs* construct	Janssen and Gardner, 1989
pNNPT	Ap^R	this study, Section 3.1.1
pUCK1	Ap^R	this study, Section 3.1.2
pUCK2	Ap^R	this study, Section 3.1.2
pDK1	Ap^R	this study, Section 3.7.1
pDK2	Ap^R	this study, Section 3.7.1
pDK3	Ap^R	this study, Section 3.7.1
pDK4	Ap^R	this study, Section 3.7.1
pDK5	Ap^R	this study, Section 3.7.1
pDK6	Ap^R	this study, Section 3.7.3

2.1.4 Antibiotics

Cefotaxime, ampicillin, carbenicillin and kanamycin were prepared as 100 mg ml^{-1} stocks in water, filter sterilised and stored at -20°C. Tetracycline was prepared as a 5 mg ml^{-1} stock ready in ethanol. When these antibiotics were incorporated into solid media, they were added just prior to pouring the petri dishes.

2.1.5 Plasmid transfer into *Agrobacterium*

Plasmids were crossed from *E.coli* into *Agrobacterium* using transfer functions supplied in trans, either as chromosomal genes in strain S17-1 (biparental cross, Simon *et al.*, 1983) or located on pRK2013 (triparental cross, Ditta *et al.*, 1980). Recipients were selected on AB minimal media which allows only *Agrobacterium* growth and on an antibiotic specific to the plasmid being crossed. Drops of fresh overnight bacterial cultures were mixed on a LB plate (40 µl *Agrobacterium*, 25 µl *E.coli*) and grown for a further 24 hours at 26°C. Then a sterile loop of cells was taken and plated for single colony growth on AB media with appropriate antibiotic. Control 'crosses' of *Agrobacterium* only or *E.coli* only were used to confirm that the original parents could not grow on the selective media. The presence of the plasmid in the new host was confirmed by plasmid DNA preparation and diagnostic restriction digestion.

2.1.6 Preparation of $CaCl_2$ competent cells

E.coli cells were made competent for naked plasmid DNA transformation by growing to the middle of their logarithmic growth phase, then by concentrating the cells and washing in $CaCl_2$ solution.

A prewarmed 40 ml broth of LB in a 250 ml erlenmeyer flask was inoculated with 1 ml of *E.coli* from an overnight culture. The flask was incubated shaking at 200 rpm until the absorbance at 600 nm reached 0.5 (approximately 2.5 hours). The culture was then chilled on ice and kept cold throughout the rest of the procedure. The cells were harvested by centrifuging at 3000 ***g*** for 10 minutes and resuspended in 20 ml of 50 mM $CaCl_2$, 15% (v/v) glycerol solution using a wide bore pipette. The cells were left on ice for 30 minutes, then centrifuged at 3000 ***g*** for 10 minutes and the supernatant discarded. Finally the cells were resuspended in 4 ml of 50 mM $CaCl_2$,

15% (v/v) glycerol solution. Aliquots of 0.3 ml could be used immediately or frozen at -70°C for later use.

2.1.7 Preparation of electrocompetent cells

E.coli cells were made electrocompetent by successive washes to remove electrolytes and concentrate the cells following the protocol of Biorad Laboratories (1989).
A 500 ml broth of LB was inoculated with 5 ml of a fresh overnight culture and incubated shaking at 200 rpm until they reached an absorbance of 1.0 at 600 nm. Then the broth was cooled on ice for 30 minutes before harvesting the cells by centrifuging at 4000 ***g*** for 15 minutes. The cells were resuspended in 500 ml of cold water, reharvested as before, resuspended in 20 ml of cold 10% (v/v) glycerol solution (not autoclaved, glycerol used directly from the suppliers bottle), reharvested a final time and resuspended in 3 ml of cold 10% (v/v) glycerol. Aliquots were stored frozen at -70°C.

2.2 Laboratory precautions

2.2.1 Government regulations

New Zealand government regulations (Advisory committee on novel genetic techniques, 1982, 1994) dealing with recombinant DNA and the transformation of plants were heeded.

2.2.2 Sterilisation

Autoclaving under steam pressure equivalent to 121°C for 15 minutes was the standard sterilisation procedure for chemical solutions, most plasticware and plant tissue culture and bacterial media, as well as being the rule for discarded bacterial and plant material, including glasshouse grown plants and their potting mix. Glassware, blotting paper and micropipette tips were autoclaved at 121°C for 45 minutes. Solutions unable to tolerate sterilisation with heat were filter sterilised through a 0.45 μm Millipore filter itself presterilised at 121°C for 15 minutes.

2.2.3 DNA precautions

All solutions, glassware and plasticware used for DNA work were sterilised to prevent contamination by micro-organisms and unwanted enzyme activity and to reduce the risk of contamination by other nucleic acids.

2.2.4 Laminar flow hood use

All plant tissue culture work and most bacterial subculturing were conducted within laminar flow hoods. These laminar flows were regularly inspected to meet Australian standards for airflow quality and pressure.
Laminar flows were prepared subsequent to use by leaving running for 15 minutes and swabbing with 95% ethanol. Forceps, blades etc were sterilised by immersing in 95% ethanol and flaming.

2.2.5 Water quality

Two grades of water were used in the laboratory. Reverse osmosis water was used for rinsing glassware after washing with detergent (pyroneg). This water was further purified through a Milli-Q filtration system (Millipore) to a resistance of 15 M ohm cm^{-1}. This water was used for all media and solutions.

2.3 DNA manipulation

2.3.1 Preparation of DNase free RNase

Pancreatic RNase (RNase A), prepared at 10 mg ml^{-1} in 15 mM NaCl, 10 mM Tris-HCl pH 7.5, was heated in a boiling water bath for 15 minutes and then left to cool slowly to room temperature before storing in aliquots at -20°C.

2.3.2 Plasmid DNA extraction method one

This procedure is derived from Holmes and Quigley's (1981) rapid boil plasmid extraction protocol. It yielded small quantities of DNA suitable for screening and

cloning.

An overnight culture of bacteria was grown shaking in 5 ml LB broth. Cells from 1.4 ml of culture were collected by 1 minute microfugation and resuspended in 350 µl of HQ-STET buffer (8% w/v sucrose, 0.05 M EDTA, 0.05 M Tris-HCl pH 8.0, 5% v/v Triton X-100). Then 25 µl of freshly prepared lysozyme solution (10 mg ml^{-1}) was added followed by gentle mixing. The tube was left standing 5 minutes before floating in a boiling water bath for 40 seconds. After 10 minutes microfugation the gelatinous pellet was discarded and the DNA precipitated with an equal volume of isopropanol (approximately 350 µl). The DNA was collected by 5 minutes microfugation, then washed with 1 ml of 95% ethanol and vacuum dried. The DNA was resuspended in 50 µl of water.

2.3.3 Plasmid DNA extraction method two

This procedure, also derived from Holmes and Quigley's (1981) rapid boil plasmid extraction protocol, is larger scale than that in Section 2.3.2 and generated purer plasmid DNA. This DNA was also suitable for double stranded DNA sequencing.

A 10 ml overnight culture grown under selection was harvested at 3000 ***g*** for 5 minutes. The cells were resuspended in 100 µl LB, then mixed with 700 µl HQ-STET buffer (see 2.3.2). Fresh lysozyme (25 µl at 10 mg ml^{-1}) was added, the tube mixed gently and left at room temperature for 10 minutes. Then the tube was floated in a boiling water bath for 2 minutes and immediately after microfuged for 10 minutes. The gelatinous precipitate was removed and 10 µl of RNase (10 mg ml^{-1}) added and mixed. After 30 minutes at 37°C two phenol/chloroform extractions were carried out and the DNA in the final aqueous phase was precipitated with an equal volume of isopropanol and microfuged for 10 minutes. The supernatant was discarded and the pellets dried, then resuspended in 100 µl of 100 mM NaCl. The DNA was then ethanol precipitated and resuspended in a final volume of 40 µl water.

2.3.4 Rapid plasmid size determination

This procedure enabled rapid size screening of newly generated plasmid clones. It is a quick alkaline lysis scaled down from a Promega protocol (Promega Corporation, 1991).

A smear of bacterial cells from a single colony was placed in a microfuge tube, then resuspended in 15 µl of 10 mM EDTA pH 8.0. Then 15 µl of cracking buffer (0.2 M NaOH, 0.5% w/v SDS, 20% w/v sucrose) was added and the tube incubated at 70°C for 5 minutes and then allowed to cool to room temperature. Then 0.5 µl of 4 M KCl and 0.5 µl of 0.4% bromophenol blue dye were added and mixed. The tube was placed on ice for 5 minutes, microfuged for 3 minutes, then a sample was analysed on an agarose minigel.

2.3.5 DNA purification by phenol extraction

DNA was extracted with an equal volume of (25:24:1 v/v) phenol:chloroform:isoamyl alcohol mixture to remove proteins, including enzymes after DNA manipulation. After mixing the upper aqueous phase containing the DNA was separated from the phenol/chloroform by a 3 minute centrifugation. Usually only one extraction was required. The DNA was then ethanol precipitated to remove traces of phenol and to concentrate the DNA.

2.3.6 Ethanol precipitation

In order to concentrate DNA and also as part of DNA purification, DNA in aqueous solution was ethanol precipitated. One tenth volume of 3 M sodium acetate, pH 5.5, was first mixed with the DNA solution, followed by 2 to 2.5 volumes of 95% ethanol, then microfugation for between 10 and 60 minutes. Longer microfugation times were used when a low concentration of DNA or small size fragments were being precipitated. The DNA pellet was washed with 70% ethanol and then 95% ethanol before either air drying or vacuum drying. The DNA was then resuspended in water.

2.3.7 Restriction endonuclease digestion

Restriction endonuclease digests of DNA were conducted with commercial preparations of enzymes and buffers. Either the manufacturers recommended buffer or the appropriate buffer from among low, medium and high salt buffers (Maniatis *et al.*, 1982) as supplied by Boehringer Mannheim was used. When necessary digests with more than one enzyme were conducted serially, digesting first with the enzyme

with lower salt requirement, then increasing the NaCl and Tris concentration to optimise conditions for the second enzyme digestion. Digests were usually for 1 to 3 hours, though plant genomic DNA was cut for longer periods, from 6 to 36 hours. Digests were incubated at the temperature (usually 37°C) recommended by the enzyme manufacturer. Digests were terminated by adding a tenth volume of SUDS (see 2.3.13) or by phenol/chloroform extraction.

2.3.8 Alkaline phosphatase treatment of DNA

In order to avoid self ligation of restriction digested vector DNA in critical cloning steps, the vector DNA was phosphatase treated to remove 3′ phosphate groups. Several units of calf alkaline phosphatase (Boehringer Mannheim) were added for the final 20 minutes of the restriction digest. The digest was then terminated by adding 20% (w/v) SDS and 0.5 M EDTA pH 8.0 to obtain final concentrations of 0.5% (w/v) SDS and 25 mM EDTA, plus 10 minutes incubation at 65°C, prior to phenol/chloroform extraction.

2.3.9 DNA ligation

Plasmid and insert DNA fragments were ligated in a volume of 10 µl with one unit of T4 DNA ligase, 2 µl of 5x ligase buffer (0.05 M $MgCl_2$, 0.25 M Tris-HCl pH 7.5, 5 mM dithiothreitol, 5 mM adenosine triphosphate). Ligation conditions usually sought were a molar ratio of 1 to 1 vector DNA to insert with a total DNA concentration of 200 ng. Ligation reactions were conducted for one hour at room temperature, or for one hour at room temperature followed by overnight at 15°C.

2.3.10 DNA transformation into $CaCl_2$ competent bacterial cells

DNA from a ligation reaction (usually 5 µl) was mixed with 0.3 ml of competent cells and left on ice for up to one hour before a heat shock of 42°C for 2.5 minutes. The cells were used to inoculate a 5 ml LB broth and grown at 37°C for 1 hour before harvesting by centrifugation at 3000 ***g*** for 5 minutes. The cells were resuspended in 200 µl LB and spread plated onto LB plates including appropriate antibiotics for selection. Plates were incubated overnight at 37°C and single colonies picked for

analysis the following day.

When necessary for blue/white selection, plates were supplemented to a final concentration of 32 μg ml^{-1} 5-bromo-4-chloro-3-indolyl β-D-galactoside (prepared as 20 mg ml^{-1} stock dissolved in dimethyl formamide, stored -20°C) and 12 μg ml^{-1} isopropyl β-D-thiogalactoside (prepared as 20 mg ml^{-1} stock, filter sterilised and stored frozen at -20°C) immediately before pouring.

2.3.11 Electroporation

Both Biorad and BRL apparatus were used for electroporation. DNA, usually 1 μl was premixed with electrocompetent cells and left on ice for 1 minute. This mix was transferred to an electroporation cuvette and pulsed (Biorad: Gene Pulser and Pulse Controller, 40 μl cells, 2 mm electrode gap, 12.5 kV cm^{-1}; settings 25 μF, 2.5 kV, 200 ohms; BRL: Cell-Porator with Voltage Booster, 30 μl cells, 1.5 mm electrode gap, 16 kV cm^{-1}; settings 330 μF, low impedance, 400 V, Voltage Booster 4000 ohms). As quickly as possible after delivering the pulse the cells were diluted in 1 ml of SOC, then grown for 1 hour at 37°C before plating on SOC plates, under appropriate antibiotic selection. Carbenicillin (100 mg ml^{-1}) was used instead of ampicillin to give reduced numbers of satellite colonies. Plates were incubated overnight at 37°C and single colonies picked for further analysis.

2.3.12 Agarose gel electrophoresis

DNA fragments were separated by electrophoresis through flat agarose gels of between 0.5 and 2% (w/v) agarose (ultrapure electrophoresis grade: GIBCO BRL Life Technologies Ltd) dissolved in Tris-acetate buffer (0.04 M Tris-acetate, 1 mM EDTA). The Tris-acetate buffer was made as a 50x stock by dissolving 121 g Tris base with 50 ml of 0.5 M EDTA pH 8.0 and 28.5 ml of concentrated acetic acid in a final volume of 500 ml. It was also used as the running buffer. Samples were loaded into wells after mixing with one tenth volume of SUDS (50% v/v glycerol, 0.1 M EDTA pH 8.0, 1% w/v SDS, 0.05% w/v bromophenol blue dye). Samples were analysed by electrophoresis at either 10 V cm^{-1} for 90 minutes (minigels) or 1 to 2 V cm^{-1} overnight (maxigels). Ethidium buffer was added to the agarose gel (and running buffer of minigels) at a concentration of 1 μg ml^{-1} in order to stain the DNA. DNA

bands were photographed on a short wavelength ultraviolet transilluminator using polaroid black and white film 667. When DNA bands were excised for further manipulation they were visualised using a handheld long wavelength (366 nm) ultraviolet source.

Molecular weight markers were obtained by digesting λ DNA with *Hin*dIII to give discernible fragments of 23.1, 9.4, 6.6, 4.4, 2.3, 2.0 and 0.56 kb, or from a *Hin*dIII, *Eco*RI double digest with fragments of 21.2, doublet of 5.1 and 4.9, 4.3, 3.5, 2.0, 1.9, 1.6, 1.4, 0.95, 0.83, and 0.56 kb, or by mixing these two combinations. Low molecular weight markers were obtained by digesting pUC8 DNA with *Hae*III to obtain DNA fragments of 587, 458, 434, 298, 267, 236, 174, 102, 80, 18 and 11 base pairs.

2.3.13 Elution of DNA fragments from agarose

Four different procedures were all used successfully to purify DNA from agarose gels after separation by gel electrophoresis.

Initially, DNA was separated by electrophoresis in a Sea Plaque (FMC Bioproducts, Maine) low melting temperature agarose gel, then DNA bands were excised in blocks of agarose, melted at 65°C and then an equal volume of TE was added, then phenol extracted and ethanol precipitated.

Later, DNA bands were electrophoresed onto DEAE-cellulose paper (Dretzen *et al.*, 1981) or into 3MM paper (Errington, 1990) inserted into a cut in the agarose below the band of interest. DNA was liberated from the DEAE-cellulose paper by treating with a high salt buffer (1.5 M NaCl, 1 mM EDTA, 20 mM Tris-HCl, pH 7.5) at 65°C for 30 minutes, followed by centrifugation to remove the paper residue. The DNA was then ethanol precipitated. DNA was eluted from 3MM paper by centrifugation through a hole cut in the base of a 500 µl tube into a 1.5 ml microfuge tube. This DNA could be ethanol precipitated, or used directly for labelling.

Alternatively, a commercial glassmilk kit ('Geneclean' from Bio 101 Incorporated or Promega Corporation 'Wizard' miniprep) was used to purify DNA from agarose slices containing the band of interest according to the manufacturers' instructions. The agarose gel was dissolved in NaI solution at 50°C, then mixed with glassmilk. The glassmilk-DNA was separated as a precipitate, then washed with a proprietary ethanol salt solution, and the DNA finally eluted into water.

2.3.14 DNA quantification

DNA was quantified using three techniques, depending on the purity and volume of the DNA solution.

Large quantities of relatively pure DNA could be quantified by measuring the absorbance at 260 nm. An absorbance of 1.0 corresponds to a DNA concentration of 50 μg ml^{-1}.

Smaller quantities of DNA could be run into a minigel and compared against serial dilutions of λ DNA to give a rough estimate.

Large numbers of samples could be reasonably accurately determined by measuring the fluorescence with Hoescht dye on microtitre plates (Rago *et al.*, 1990). Samples were prepared in triplicate with 20 μl of DNA solution (diluted as necessary), 80 μl buffer (2 M NaCl, 1 mM EDTA, 10 mM Tris-HCl pH 7.4) and 100 μl Hoescht dye (20 μg ml^{-1} in aforementioned buffer, diluted from 50x stock of 1 mg ml^{-1} dye in water stored at 4°C). Standards of λ DNA of 50, 40, 30, 20, 10 and 0 μg ml^{-1} were used. Samples were measured in an automated 96 well fluorescent plate reader (Perkin Elmer LS50B) using an excitation wavelength of 350 nm and emission of 460 nm.

2.3.15 DNA sequencing

DNA sequencing was performed using the United States Biochemical company's 'Sequenase' kit for dideoxy chain termination with a double stranded DNA template, incorporating [α-^{35}S] deoxyadenosine-5′-triphosphate nucleotides (1000 Ci $mmol^{-1}$, 10 μM, ICN Pharmaceuticals), followed by polyacrylamide gel electrophoresis and autoradiography. 'Sequenase' is a recombinant T7 DNA polymerase without 3′-5′exonuclease activity.

'Sequenase' kit instructions (United States Biochemical, 1992) were followed precisely for DNA denaturation, annealing to primer, chain extension and termination and stopping the reactions.

A standard 6% polyacrylamide gel solution was prepared by mixing 15 ml of 40% (w/v) acrylamide stock (acrylamide:bisacrylamide, 19:1), 50 g urea, 10 ml 10x TBE buffer (121 g l^{-1} Tris base, 55 g l^{-1} boric acid, 7.4 g l^{-1} Na_2EDTA, pH 8.3) and 35 ml water until the urea was dissolved. Then 1 ml of 10% (w/v) ammonium

persulphate and 20 µl NNN'N'-tetramethylethylenediamine were added and the solution poured between 0.4 mm spaced plates made to fit GIBCO BRL model S2 sequencing apparatus; and a sharks tooth comb was added.

The gel was run with 1x TBE buffer at a constant power of 65 Watts. Typically 2 sets of samples, using 3 µl reaction mix per lane, were run; one set run for 2 hours and another for 5 hours to give a longer readable sequence. The gel was transferred to 3MM paper, vacuum dried and autoradiographed using Fuji RX film.

2.4 Tobacco growth

2.4.1 Tissue culture growth room

Tobacco plants in tissue culture were grown at a controlled temperature of 26°C in a growth room under 24 hour light supplied by a bank of 8 cool white fluorescent tubes (58 watts each) offset in the ceiling.

2.4.2 Tissue culture media

A variety of growth containers were used for plant tissue culture, from standard petri dishes (85 mm diameter, 13 mm high, 25 ml media) for seed germination and *Agrobacterium* inoculation of leaf sections, to deep petri dishes (same diameter, 23 mm high, 30 ml media) for regeneration and pottles (90 mm diameter, 55 mm high, 40 ml media) for individual rooted plants. All media (except water agar) was adjusted to pH 5.7 with KOH and was solidified with 0.8% (w/v) Difco Bacto agar. Naphthalene acetic acid and 6-benzylamino-purine were prepared as 1 mg ml^{-1} stocks, first dissolved in 1 M NaOH then made to volume. Stock solution of B5 vitamins (1 g l^{-1} nicotinic acid, 10 g l^{-1} thiamine HCl, 1 g l^{-1} pyridoxine HCl; Gamborg *et al.*, 1968) was stored frozen in 1 ml aliquots.

Stock solutions A to F (salts) and G (vitamins) were prepared for MS based media, based on the formula of Murashige and Skoog (1962). For one litre of media, 20 ml of stocks A (82.5 g l^{-1} NH_4NO_3) and B (95 g l^{-1} KNO_3), 5 ml of stocks C (1.24 g l^{-1} H_3BO_3, 34 g l^{-1} KH_2PO_4, 0.165 g l^{-1} KI, 0.05 g l^{-1} $Na_2MoO_4.2H_2O$, 0.005 g l^{-1} $CoCl_2.6H_2O$), D (88 g l^{-1} $CaCl_2.2H_2O$), E (74 g l^{-1} $MgSO_4.7H_2O$, 4.46 g l^{-1} $MnSO_4.4H_2O$, 1.71 g l^{-1} $ZnSO_4.7H_2O$, 0.005 g l^{-1} $CuSO_4.5H_2O$), F (8 g l^{-1} ferric

EDTA) and when needed 5 ml of vitamin stock G (0.1 g l^{-1} nicotinic acid, 0.02 g l^{-1} thiamine HCl, 0.1 g l^{-1} pyridoxine HCl, 0.4 g l^{-1} glycine) were used. An additional 100 mg l^{-1} of myo-inositol was added as part of the MS salts' requirement.

water agar	0.8% (w/v) agar only.
MS	MS salts and vitamins with 30 g l^{-1} sucrose added.
1/2 MS	Half concentration of MS salts and vitamins with 15 g l^{-1} sucrose.
NicI	MS salts with 1 ml of B5 vitamin stock per litre and 30 g l^{-1} sucrose plus 1 mg l^{-1} 6-benzylamino-purine and 0.1 mg l^{-1} naphthalene acetic acid. No antibiotics were added. (Horsch *et al.*, 1985).
NicII	NicI media usually with both kanamycin (100 or 300 mg l^{-1}) and cefotaxime (100 or 500 mg l^{-1}).
NicIII	NicII media without 6-benzylamino-purine and naphthalene acetic acid.

2.4.3 Tobacco genotypes

Two *Nicotiana tabacum* genotypes were used, Wisconsin 38 supplied courtesy of Dr TJ Higgins (CSIRO, Canberra) and in house DSIR line KKD.

2.4.4 Seed sterilisation

Tobacco seed were surface sterilised for growth in tissue culture. Several hundred seed were tied into a small muslin bag with a metal 'plastic-bag' tie. The bag of seed was soaked in 25% (v/v) commercial strength hypochlorite solution with a drop of detergent (1% v/v cetavalon) for 20 minutes with regular agitation. Then the bag was rinsed five times with water and the seed plated onto either water agar or 1/2 MS media.

2.4.5 *Agrobacterium*-mediated tobacco transformation and regeneration

Transformation of *Agrobacterium* followed the protocol of Horsch *et al.* (1985) using a binary plasmid system (Hoekema *et al.*, 1983). *Agrobacterium* cultures were grown 24 hours at 26°C in 5 ml TY broth with shaking under appropriate antibiotic selection to maintain the T-DNA containing plasmid (Koncz and Schell, 1986; pPCV series

vectors without selection segregate in *Agrobacterium* at a rate of 2.5% per generation). Cells were collected by centrifugation at 3000 ***g*** for 10 minutes and resuspended in 10 ml of 10 mM $MgSO_4$. Tobacco leaf sections of 0.4 to 1.0 cm^2, avoiding the midrib, were cut from axenic tobacco plants and immersed in *Agrobacterium* solution for less than 10 seconds. The leaf sections were then blotted dry on both sides and placed with the lower epidermis up onto NicI media. Two to four days later the leaf sections were transferred to NicII media with both kanamycin to kill untransformed plant cells and cefotaxime to inhibit *Agrobacterium* growth. Four to six weeks later, usually after one transfer onto fresh NicII media, individual shoots or leaf sections developing shoots were transferred to NicIII media with the same antibiotic selection. Individual shoots were rooted on NicIII media in pottle to form individual plants. To maintain good health, fresh shoot cuttings were taken every 6 weeks and rerooted on NicIII.

2.4.6 Glasshouse tobacco growth

Rooted plants in tissue culture were transferred to commercial potting mix (Midland Horticulture, 60% peat, 40% sand supplemented with a slow release fertiliser) by carefully washing the agar off the roots and setting the plants in soaked potting mix in 10 cm deep, 12 cm diameter plastic planter bags. Plants were kept covered to maintain a high humidity until they were established. The glasshouse had cooling in summer and basic heating in winter to maintain the temperature in the range 10 to 25°C.

Plants were occasionally sprayed with 5 ml l^{-1} Target (Yates) and 1 g l^{-1} Pirimor (ICI) with surfactant, 1 ml l^{-1} Sprayfix (Yates), to control whitefly and aphids.

Plants were allowed to self-pollinate to produce seed. Control flowers were emasculated by removing pollen sacs from immature flowers 1 to 2 cm long with sterile forceps. These flowers never set seed unless pollen was transferred onto their stigma by hand.

2.4.7 Analysis of progeny

To test for segregation of kanamycin resistance, seedlings were germinated on 1/2 MS or water agar containing 200 mg l^{-1} kanamycin (Paszkowski and Saul, 1986).

To decrease methylation of cytidine within DNA, seedlings were germinated on media containing 5-azacytidine at 50, 30 or 5 μM, prepared as a filter sterilised stock of 40 mM, stored at -20°C.

2.5 Analysis of β-glucuronidase activity

2.5.1 Standard histochemical assay

Plant material was stained for β-glucuronidase activity by immersing in stain solution (0.5 mg ml^{-1} XGluc, 0.1% v/v triton X-100, 0.1% v/v β-mercaptoethanol, 50 mM sodium phosphate pH 7.0) for 24 hours at 37°C. Material was cleared after staining to remove chlorophyll by soaking in 50% (v/v) ethanol for 2 hours with gentle shaking, followed by successive changes of 95% (v/v) ethanol until the green colour was removed.

XGluc was prepared as a 2x stock, stored at -20°C, at 1 mg ml^{-1} in 100 mM sodium phosphate pH 7.0 after dissolving in dimethyl formamide.

2.5.2 Floral histochemical assay

After trials to reduce oxidative browning (refer Section 4.3) the following protocol was adopted. Floral material was sectioned under a solution of 50 mM sodium phosphate, 0.1% (v/v) triton X-100, 100 mM sodium ascorbate pH 7.0 before transferring to stain solution (0.5 mg ml^{-1} XGluc, 100 mM sodium ascorbate, 0.5 mM ferricyanide, 0.5 mM ferrocyanide, 0.1% v/v triton X-100, 0.1% v/v β-mercaptoethanol, 50 mM sodium phosphate pH 7.0) and vacuum infiltrating (water tap pressure) for 10 minutes before incubation for 24 hours at 37°C. Chlorophyll was cleared after staining as described above (Section 2.5.1).

The floral material tested was: a longitudinally cut flower bud (stage -1, 1 or 2, Koltunow *et al.*, 1990); a transverse section through a seed pod and surrounding sepals; and from a just opened flower, an anther and top of the filament, the stigma and top of the style, and a longitudinal section through the ovary, receptacle, base of the style and sepals.

2.5.3 Enzyme extraction for fluorometric assay

GUS activity was determined following the procedure of Jefferson (1987).
Plant samples were deep frozen with liquid nitrogen in a microfuge tube and ground to a powder with a homemade araldite grinder (made by setting a metal handle in Araldite glue within a microfuge tube) and then reground with 200 µl of extraction buffer (50 mM sodium phosphate pH 7.0, 10 mM β-mercaptoethanol, 10 mM Na_2EDTA, 0.1% w/v sodium lauryl sarcosine, 0.1% v/v triton X-100). The extract was microfuged for 3 minutes and the supernatant used for GUS enzyme assays.
The protein concentration of the extract was determined with the Biorad protein determination kit, which uses the method of Bradford (1976). Samples containing up to 25 µg of protein were diluted in 0.8 ml of water and 0.2 ml of Biorad kit coomassie dye added. Absorbance was measured at 595 nm and compared to standards of 0 to 30 µg bovine serum gamma globulin. Samples were determined in triplicate.

2.5.4 Microtitre plate fluorometric assay

Samples were assayed for GUS activity on microtitre plates using the substrate methyl umbelliferone glucuronide (MUG). Each plant extract was assayed in three wells with 10 µl of enzyme extract (2.5.3) added to 50 µl of assay buffer (extraction buffer, see 2.5.3, with 1 mM MUG added). The three wells were stopped at three different time points, zero time (ie before enzyme extract added), 1 hour and 6 hours by adding 50 µl of 0.5 M Na_2CO_3. Plates were incubated at 37°C for 6 hours before scoring fluorescence by eye under illumination from a handheld long wavelength ultraviolet light source of 366 nm. When samples were thought to contain endogenous fluorescence, control assays without MUG were conducted.

2.5.5 Standard fluorometric assay

GUS assays were set up in duplicate. Aliquots (2 to 50 µl) with approximately 20 µg of plant protein (2.5.3) were added to 500 µl of prewarmed assay buffer (2.5.4). The assay was allowed to equilibrate (about 10 minutes) at 37°C before the first sample of 100 µl was taken and added to 900 µl of 0.2 M Na_2CO_3 to stop the reaction. Two further samples were taken at half hour intervals. Fluorescence was measured in a

fluorometer (TK0 100, Hoefer Scientific Instruments) with excitation wavelength 365 nm and emission wavelength 460 nm, after adding a further 1 ml Na_2CO_3 to make up to 2 ml. The fluorometer was zeroed with 0.2 M Na_2CO_3 solution on its most sensitive scale and then calibrated with methyl umbelliferone standards of 0 to 200 nM made up in 0.2 M Na_2CO_3. Readings were taken with the fluorometer set to 500 with 50 nM methyl umbelliferone solution.

2.6 Southern blotting and hybridisation

2.6.1 Tobacco DNA extraction

Tobacco DNA was extracted using a CTAB extraction procedure modified from Doyle and Doyle (1987) by Andrew Griffiths (personal communication).

Leaf material (0.5 - 1 g) was ground to a powder with a mortar and pestle after freezing with liquid nitrogen and then added to 7.5 ml of preheated CTAB extraction buffer (2% w/v CTAB, 1.4 M NaCl, 0.2% v/v β-mercaptoethanol, 20 mM EDTA, 100 mM Tris-HCl pH 8.0) and mixed thoroughly before incubating at 65°C for 30 to 60 minutes. Then 7.5 ml of chloroform/isoamyl alcohol (24:1 v/v) was added and mixed and the sample centrifuged for 5 minutes at 12000 ***g***. Two thirds volume of isopropanol was added to the supernatant and the nucleic acid precipitate collected by centrifugation at 3000 ***g*** for 2 minutes. The supernatant was discarded and the pellet allowed to air dry before resuspending in 1 ml of TE buffer with 0.1 mg of RNase (Section 2.3.1) and incubating at 37°C for 30 minutes. The DNA was reprecipitated by adding 2 ml water, 1.5 ml of 7.5 M ammonium acetate pH 7.7 and 10 ml ethanol and collected by centrifuging at 3000 ***g*** for 2 minutes. The pellet was washed with 80% (v/v) ethanol and air dried before resuspending in 0.5 ml water.

2.6.2 Southern blotting

DNA was blotted onto Hybond N+ nylon membrane (Amersham) following the manufacturer's instructions, after the method of Southern (1975).

Agarose gel electrophoresis was conducted with typically 20 μg of restriction digested plant genomic DNA per lane on a maxigel (Section 2.3.12). Gels were photographed, trimmed and soaked in 0.25 M HCl for 15 minutes, rinsed with water, soaked in

denaturation buffer (1.5 M NaCl, 1.5 M NaOH) for 30 minutes, then rinsed with water and soaked in two changes of neutralisation buffer (1.5 M NaCl, 1 mM EDTA, 0.5 M Tris-HCl, pH 7.2) for 15 minutes each. A wick of 3 sheets of Whatman 3MM, sitting in reservoirs of 20xSSC (3M NaCl, 0.3M Na_3 citrate) was set up, topped with 3 rectangles of 3MM cut to match the agarose gel, which was placed on top. Then a matched rectangle of Hybond N+ membrane, handled only with gloves, was placed on the agarose gel, followed by 3 sheets of size cut 3MM prewetted in 20xSSC, followed by stacks of dry paper towels and a 1 kg weight on a glass plate. After blotting upward overnight, the membrane was rinsed for 30 seconds in 2xSSC. The DNA was fixed to the membrane by placing on Whatman 3MM soaked in 0.4 M NaOH (DNA side up) for 20 minutes. Finally the membrane was rinsed in 5xSSC before sealing in a plastic bag and storing at 4°C ready for hybridisation.

2.6.3 Labelling probe

DNA, usually after gel extraction and purification, was labelled using a Hi-Prime kit (Boehringer Mannheim) or Ready To Go kit (Pharmacia Biotech) according to the manufacturers' instructions. About 25 ng of DNA dissolved in water was mixed with kit components (Klenow fragment DNA polymerase, random primers and nucleotides) and 5 µl [α-^{32}P] deoxycytidine-5′-triphosphate (3000 Ci mmol^{-1}, 3.3 µM, ICN Pharmaceuticals).

Incorporation of [α-^{32}P]-dCTP into the probe DNA was confirmed by thin layer chromatography on Cellulose F_{245}-coated aluminium (Merck) with 0.75 M KH_2PO_4, pH 3.5 (adjusted with orthophosphoric acid) followed by autoradiography. Unincorporated nucleotide moved with the solvent, while incorporated nucleotide remained at the origin.

2.6.4 Hybridisation

DNA fixed onto a membrane by Southern blotting was hybridised with ^{32}P-labelled DNA probes using a modification of the protocol of Church and Gilbert (1984) in glass tubes in a rotary Hybaid oven.

Membranes and Hybaid mesh were prewetted in 2xSSC, and the membranes placed between layers of mesh, excluding air bubbles, and rolled into a cylinder and placed

in the hybridisation glass tube. Up to five membranes could be hybridised in one tube. Once the membranes were correctly unrolled in the glass tube, the 2xSSC was replaced with 50 ml of prehybridisation solution (0.5 M sodium phosphate, pH 7.2, 1 mM EDTA, 7% w/v SDS) and rotated in the Hybaid oven at 65°C for an hour. The prehybridisation solution was decanted and replaced with 20 ml of hybridisation solution (same composition as prehybridisation solution) with the ^{32}P-labelled probe DNA added. The membranes were then incubated with rotation at 65°C overnight. Unbound probe DNA was washed off with 5 washes at 65°C in the Hybaid oven, undergoing rotation, with 100 ml preheated wash solution. The first two washes were 15 minutes each with 2xSSC including 0.1% (w/v) SDS, then 30 minutes with 1xSSC including 0.1% (w/v) SDS and the final two washes for 10 minutes each with 0.1xSSC including 0.1% (w/v) SDS. The membranes were rolled up, removed from the glass tube and individually sealed in plastic bags and autoradiographed with Fuji RX film and an intensifier screen at -70°C.

2.6.5 Stripping

Nylon membranes were stripped of radioactive label by heating gently in a microwave in 0.5% (w/v) SDS until the solution boiled, allowed to cool to room temperature, and then washed in 2xSSC and sealed in a plastic bag and kept at 4°C until the next hybridisation.

2.7 Rescue of flanking plant genomic sequence

2.7.1 Polymerase chain reaction

The polymerase chain reaction (Saiki *et al.*, 1988) was used to amplify regions of DNA between known primer sequences or, as detailed below, external of specific primer sequences. The technique was used to generate DNA for making probes for hybridisation using universal and reverse primers, for inverse PCR (2.7.3), single sided and ligation mediated PCR (2.7.5).

The standard conditions for PCR reactions were 10 µl 10x PCR buffer (supplied by manufacturer), 10 µl deoxynucleotide stock solution (2 mM each nucleotide), 1 µl of

each primer from 20 pmol l^{-1} stocks plus either 1 μg plant genomic DNA or 1 ng plasmid DNA. The reaction was made up to 100 μl with water and 0.5 μl Taq polymerase (Cetus Amplitaq or Promega Taq) added prior to thermocycling, usually for 30 cycles. Either a Perkin Elmer thermocycler or a Hybaid thermocycler were used with block temperature control, but monitored with an eppendorf tube thermocouple. Cycle conditions are detailed in the results.

Univeral and reverse primers were used with a cycle of 1 minute 94°C, 1 minute 50°C, 1 minute 72°C, for 30 cycles.

2.7.2 PCR primers

Oligonucleotide primers used for PCR, DNA sequencing and the initial DNA synthesis in single sided ligation mediated PCR are listed in Table 2.

The primers RBG, DSO, USN and DSG were selected from the appropriate DNA sequence (Figure 2) with the aid of the computer programme Primer Designer (Scientific and Educational Software) to avoid homologies which could lead to self priming or dimerisation of primer pairs. Sequences were chosen which ended in two guanidine or cytidine nucleotides.

The melting temperatures of the primers were estimated using the formula from Mueller and Wold (1990):

$$\text{Melting temperature} = 81.5 + 16.6(\log M) + 0.41(\%GC) - 500/n$$

Where M is the molarity of salt in the reaction buffer (0.051M) and n is the length of the primer.

The primer RBG has a one base pair mismatch at its extreme 5′ end and so its effective melting temperature in initial rounds of PCR will be 66°C. Otherwise all primers exactly matched the T-DNA sequence.

Primers were either purchased from Oligos Incorporated or synthesised in house by Dr Paul Ealing. When necessary oligodeoxynucleotides were purified (after deprotection) according to the protocol of Sawadogo and Van Dyke (1991). The oligodeoxynucleotide solution in 30% (v/v) NH_4OH was vortexed vigorously for 15 seconds with 10 volumes of n-butanol, then microfuged. The supernatant was discarded and the oligodeoxynucleotide containing pellet was vacuum dried and resuspended in water.

Table 2. Oligonucleotide primers.

Primer	Sequence 5' to 3'	Melting temperature[1]
RBG	GCC GGA TCC CCG GGA TCA GAT TGT CG	68°C
DSO	GCA ACT GCT TTG TTC GGT TCG GGA CC	64°C
USN	GAT AGC CGC GCT GCC TCG TCC	66°C
DSG	CAA CAA CTC TCC TGG CGC ACC ATC G	65°C
$FOOT_{11}$	GAA TTC AGA TC	29°C
$FOOT_{25}$	GCG GTG ACC CGG GAG ATC TGA ATT C	65°C
Universal	GTA AAA CGA CGG CCA GT	52°C
Reverse	AAC AGC TAT GAC CAT G	47°C

[1] calculated according to the formula given by Mueller and Wold (1990). Refer text.

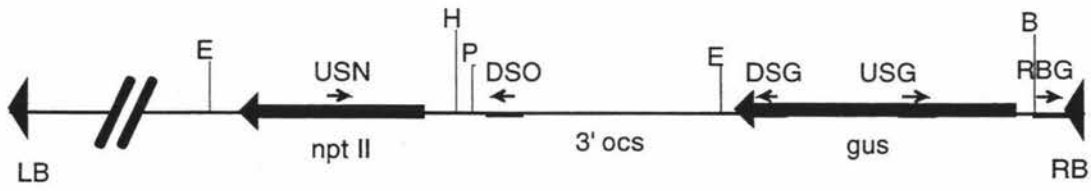


Figure 2. Primer sites within the T-DNA region of pGTG and pBin19-GTG.

The positions and orientations of the five primers USN, DSO, DSG and RBG are displayed above the DNA line (primer lengths not drawn to scale). The positions of the *gus* and *npt*II genes and the *ocs* terminator sequence are shown below the line. Restriction sites are marked with a single letter code: E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; B, *Bam*HI. Not all sites within the T-DNA are marked for these enzymes.

The concentration of the oligodeoxynucleotides was estimated from the absorbance of an aliquot at 260 nm and the extinction coefficient calculated from addition of the individual nucleotide extinction coefficients, 11.4 per guanidine residue, 15.4 per adenine residue, 9.7 per thymine residue and 9.2 per cytidine residue.

2.7.3 Inverse PCR

Sequence flanking the right border of T-DNA insertions into plant genomic DNA could be amplified for cloning using inverse PCR (Triglia *et al.*, 1988; Ochman *et al.*, 1988).

Plant genomic DNA (2.6.1) was restriction digested (2.3.7), purified by phenol extraction (2.3.5) and/or ethanol precipitation (2.3.6) and ligated (2.3.9), albeit in a much larger volume (typically 20 μg DNA in 1 ml) to encourage only self-ligation of DNA fragments. This DNA was then ethanol precipitated and usually 1 μg used per PCR with primer pairs of either RBG/DSO (following *Hin*dIII or *Pst*I digestion) or RBG/DSG (following *Eco*RI digestion). After a brief heat treatment at 94°C to nick the DNA (Triglia *et al.*, 1988), the sample was subjected to PCR. Details of cycle conditions are presented in Section 3.7.1

The USN/DSO primer pair were used to optimise PCR conditions as they amplified an internal T-DNA fragment of 0.6 kb without requiring digestion and ligation.

2.7.4 Plasmid rescue

The pBR322 β-lactamase and origin of replication within the T-DNA of the promoter tagging vectors could be 'rescued', along with flanking sequence potentially of plant origin, as a plasmid from plant genomic sequence within transgenic plants.

Plant genomic DNA (2.6.1) was restriction digested (2.3.7), purified by phenol extraction (2.3.5) and/or ethanol precipitation (2.3.6) and ligated (2.3.9), albeit in a much larger volume (typically 20 μg DNA in 1 ml) to encourage only self-ligation of DNA fragments. This DNA was then ethanol precipitated and resuspended in a volume suitable for transformation, usually by electroporation (2.3.11). For specific examples refer Section 3.7.3.

In order to avoid false positive transformants from contaminating plasmid DNA the following measures were adopted. Strict precautions were taken, isolating all

glassware, plasticware and solutions used to prepare genomic DNA and to prepare competent cells. Glassware etc was either bought new or cleaned by soaking in 10% (v/v) hypochlorite solution overnight to remove nucleic acids and thereafter kept out of communal cleaning arrangements.

2.7.5 Single sided ligation mediated PCR

Plant genomic DNA flanking a T-DNA insertion could be amplified with a single T-DNA primer facing externally by single sided ligation mediated PCR closely following the protocol of Fors *et al.* (1990) with additional explanation provided by unpublished notes (Mueller and Wold, 1990) and an earlier publication (Mueller and Wold, 1989). The RBG primer (0.3 pmol) was annealed to 3 μg of restriction digested DNA in 15 μl of 40 mM Tris-HCl (pH 7.7), 50 mM NaCl by heating at 95°C for 2 minutes, then at 60°C for 30 minutes. The mix was placed on ice and 7.5 μl of 20 mM $MgCl_2$, 20 mM dithiothreitol, 0.02 mM each deoxynucleotide-triphosphate was added followed by 1.5 μl of a 1:4 dilution of Sequenase (US Biochemicals) diluted in 10 mM Tris-HCl (pH 7.5), 1 mM EDTA. Primer extension was carried out at 47°C for 5 minutes and the reaction stopped by heating at 67°C for 10 minutes after adding 6 μl of 310 mM Tris-HCl (pH 7.7) and then placed on ice. Next 20 μl of 17.5 $MgCl_2$, 42.3 mM dithiothreitol, 125 μg ml^{-1} bovine serum albumin was added followed by 25 μl of ligation mix bringing the solution to 10 mM $MgCl_2$, 20 mM dithiothreitol, 3 mM adenosine triphosphate, 50 μg ml^{-1} bovine serum albumin, 50 mM Tris-HCl (pH 7.7), 100 pM linker and 3 units of T4 DNA ligase. Linker was prepared by denaturing 20 pmol μl^{-1} each of $Foot_{11}$ and $Foot_{25}$ primers in 250 mM Tris-HCl (pH 7.7) at 95°C for 5 minutes, then annealing by cooling from 70°C to room temperature slowly (one hour), leaving at room temperature for one hour, gradually cooling to 4°C and leaving at 4°C for 24 hours before storing at -20°C.

The ligation was incubated overnight at 19.5°C, ethanol precipitated and then subjected to PCR using RBG and $Foot_{25}$ primers for 30 or 40 cycles of 94°C for 1 minute, 66°C for 1 minute, 72°C for 3 minutes, ultimately followed by 10 minutes at 72°C. An aliquot was examined by gel electrophoresis.

2.7.6 Analysis of DNA sequence

DNA sequences were compared with the Genbank database at the National Centre for Biotechnology Information, National Institute of Health, USA using electronic mail (blast@ncbi.nlm.nih.gov) and the alignment programme Blastn (Altschul *et al.*, 1990).

3.0 Results I

3.1 Promoter tagging vector construction

The promoter tagging vectors pGTG and pBin19-GTG were constructed starting from pPCV604 (Koncz, 1989; Figure 1). An overview of the cloning steps is presented in Figure 3. In summary, a partial digest of pPCV604 removed 100 base pairs and allowed a fragment from pBin6 (Bevan, 1984) coding for kanamycin resistance to be added, creating pGT. Then a fragment of pKIWI101a (Janssen and Gardner, 1989) was subcloned and ligated into pGT, replacing the promoterless *hpt* gene with a promoterless *gus* gene, creating pGTG. Finally, the vector sequence outside the T-DNA of pGTG, derived from pPCV604, was replaced with the vector functions of pBin19 (Bevan, 1984) creating pBin19-GTG.

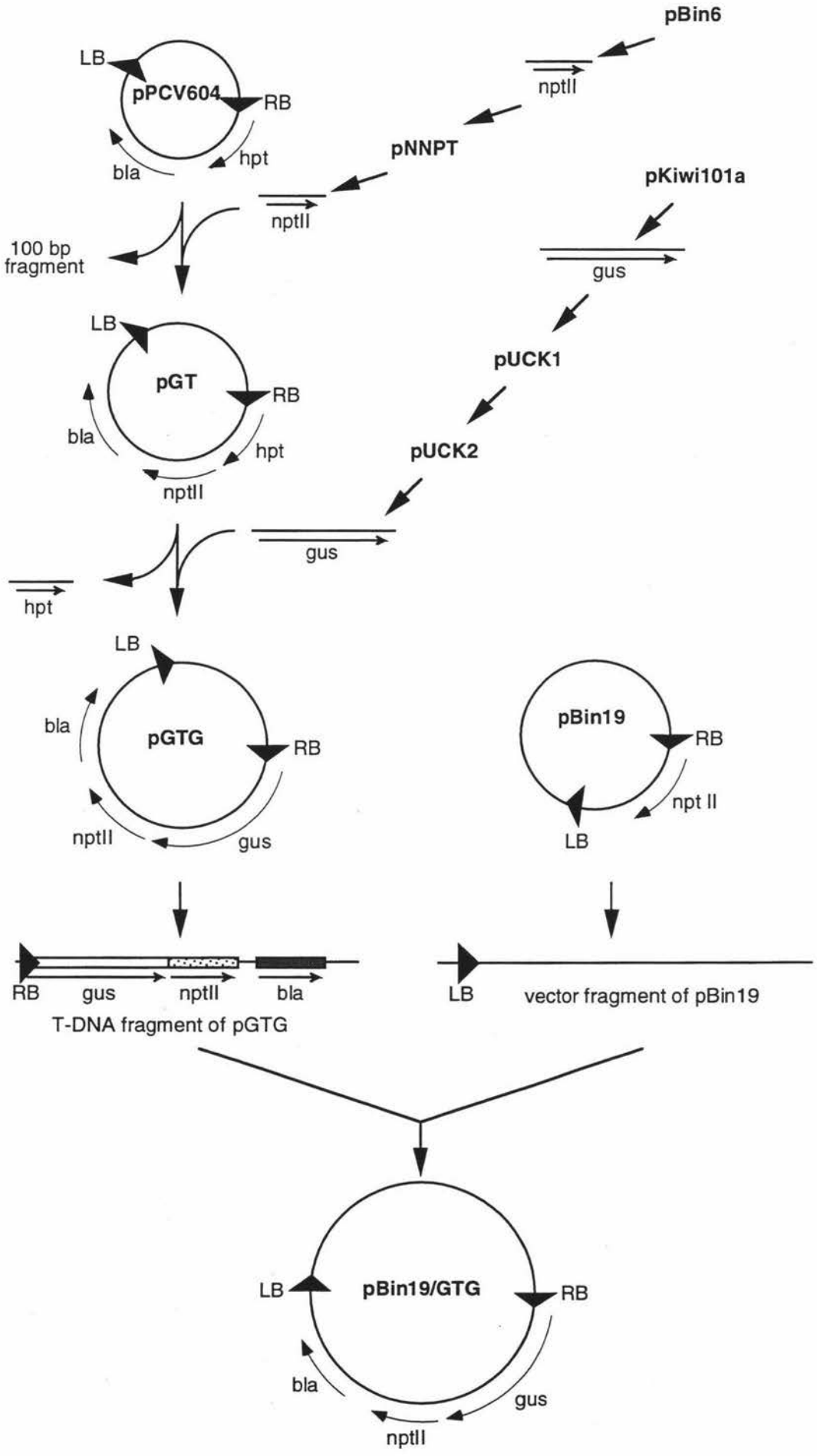
3.1.1 Construction of pGT

The plasmid pBin6 (Appendix 8.1.1) was available in *Agrobacterium* strain LBA4404. In order to increase plasmid yields for subcloning, plasmid DNA was prepared from LBA4404 (Section 2.3.2) and transformed into *E. coli* strain S17-1.

Interpreting Bevan's (1984) description, pBin6 was thought to have a unique restriction site for *Hin*dIII and two sites for *Eco*RI, with the plant selectable kanamycin resistance gene residing on a 1.6 kb *Hin*dIII-*Eco*RI fragment. This gene consists of the *npt*II coding sequence and *nos* 5′ promoter and 3′ processing signals. However *Hin*dIII and *Eco*RI digests revealed two sites for each enzyme. To substantiate that the appropriate plasmid was being used, it was restriction mapped with *Hin*dIII, *Eco*RI and *Bgl*II. The third enzyme, *Bgl*II, was used to reduce fragment sizes for greater accuracy. The restriction fragments sizes created with single enzyme digests and each combination of double digests (Figure 4) were estimated (Table 3). The mapped plasmid (Appendix 8.1.1) had a size of 19 kilobases, which is 3 to 4 kb greater than the two sizes, 15 and 16 kb, both reported by Bevan (1984). However, 1.5 kb *Eco*RI-*Eco*RI (Figure 4, lanes 3 and 8) and 1.6 kb *Hin*dIII-*Eco*RI (Figure 4, lane 3) fragments, were found adjacent (Appendix 8.1.1), as expected, so the cloning

Figure 3. Overview of pBin19-GTG plasmid construction.

The details of the plasmid construction are recorded in the text (Section 3.1)
Abbreviations: LB and RB, left and right borders of T-DNA; *hpt*, promoterless hygromycin phosphotransferase gene; *gus*, promoterless β-glucuronidase gene; *npt*II, chimaeric neomycin phosphotransferase gene with nopaline synthase promoter and 3′ processing signals; *bla*, β-lactamase gene.


pBin6
nptII
LB
pPCV604
RB
hpt
bla
pNNPT
nptII
pKiwi101a
100 bp
fragment
gus
LB
pGT
RB
bla
hpt
nptII
pUCK1
pUCK2
gus
hpt
LB
pGTG
RB
bla
nptII
gus
pBin19
RB
npt II
LB
RB
gus
nptII
bla
T-DNA fragment of pGTG
LB
vector fragment of pBin19
LB
pBin19/GTG
RB
bla
gus
nptII

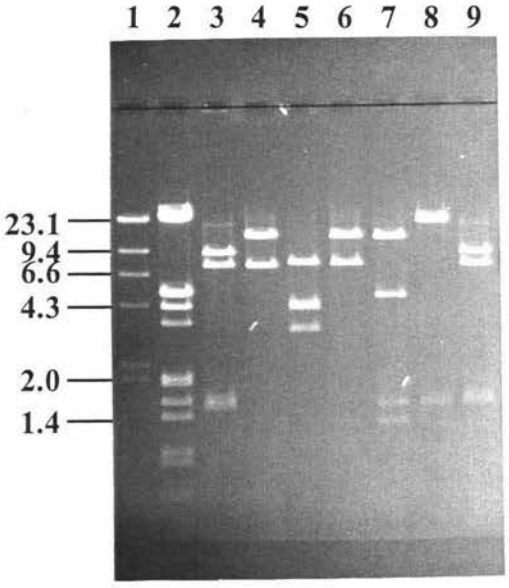


Figure 4. Restriction endonuclease digestion of pBin6.

Lane 1: *Hin*dIII digested λ DNA size markers (Section 2.3.12). Lane 2: *Hin*dIII, *Eco*RI double digested λ DNA size markers (Section 2.3.12). Remaining lanes are pBin6 DNA digested with *Hin*dIII and *Eco*RI (lane 3), *Hin*dIII (lane 4), *Hin*dIII and *Bgl*II (lane 5), *Bgl*II (lane 6), *Bgl*II and *Eco*RI (lane 7), *Eco*RI (lane 8) and *Hin*dIII and *Eco*RI (lane 9). Numbers on the left of the figure indicate band sizes in kb.

Table 3. Fragment sizes of pBin6 as determined from Figure 4.

Restriction digest	Fragment size (kb)
*Eco*RI	17.5, 1.5
*Hin*dIII	11.8, 7.2
*Bgl*II	11.5, 7.5
*Eco*RI/*Hin*dIII	8.7, 7.2, 1.6, 1.5
*Eco*RI/*Bgl*II	11.5, 4.7, 1.5, 1.3
*Hin*dIII/*Bgl*II	7.4, 4.4, 4.1, 3.1

was continued.

The plasmid Bin6 (Appendix 8.1.1) was double digested with *Hin*dIII and *Eco*RI. The sample was split into two and either phenol extracted and ethanol precipitated (Sections 2.3.5 and 2.3.6), or run on a Sea Plaque agarose gel and the lower bands (1.5 and 1.6 kb, eg Figure 4, lane 3) cut out of the gel and extracted with phenol (Section 2.3.13). The vector pUC8 (Appendix 8.1.2) was prepared with *Hin*dIII and *Eco*RI double digestion and then split into two samples. To prevent religation with the small polylinker fragment, the vector was either alkaline phosphatase treated (Section 2.3.8) or run on a Sea Plaque agarose gel and the linear vector band extracted (Section 2.3.13). Ligations (Section 2.3.9) were set up for all four combinations of vector and pBin6 fragments and the DNA was transformed (Section 2.3.10) into S17-1. All four transformations resulted in colonies able to grown on LB 50 mg l^{-1} ampicillin. The combination of Sea Plaque extraction of both vector and insert resulted in a quarter the number of transformants compared to the other three treatments. Control transformations with vector only from both preparations resulted in zero transformants. Eight transformants were screened with *Hin*dIII and *Eco*RI single and double digests. Seven colonies had insertions of the 1.6 kb fragment and one colony had an insertion of an 8.6 kb *Hin*dIII-*Eco*RI fragment. One of the transformants with a smaller insertion was selected and the new plasmid named pNNPT (Figure 5).

The plasmid pNNPT (Figure 5) was digested with *Hin*dIII and *Eco*RI and the 1.6 kb fragment was phenol extracted from a Sea Plaque gel (Section 2.3.13) prior to cloning into *Hin*dIII, partial *Eco*RI digested pPCV604 (Figure 1) as described below.

In order to insert the 1.6 kb chimaeric *npt* gene into pPCV604, it was necessary to cut the plasmid at the unique *Hin*dIII site and at the closer one of the two *Eco*RI sites, causing a small deletion of 100 base pairs (Figure 1). Partial digests of *Eco*RI were optimised for yield of the linear fraction by serial dilution of the enzyme from 1 to 0.03 units of enzyme in 45 minute digests (Figure 6). The best yield of 9 kb fragment was with 0.125 units of *Eco*RI (Figure 6, lane 6). Appropriately digested DNA was separated by electrophoresis in a Sea Plaque agarose gel (Section 2.3.13) and the 9 kb linear band was cut out of the gel, attempting as best as possible to avoid supercoiled

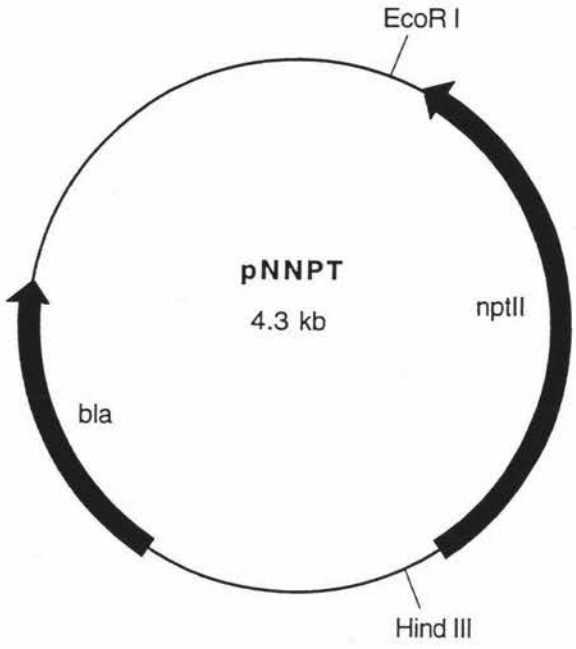


Figure 5. Physical map of pNNPT.

A 1.6 kb *Hin*dIII-*Eco*RI fragment containing the chimaeric Pnos−*npt*II−3′*nos* gene (*npt*II) was inserted into *Hin*dIII/*Eco*RI cut pUC8.

Abbreviations: *bla*, β-lactamase; *npt*II, neomycin phosphotransferase II.

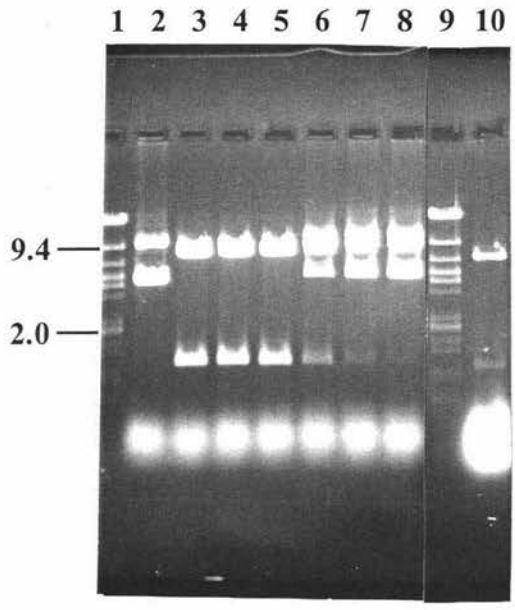


Figure 6. Partial digestion of pPCV604 with *Eco*RI.

pPCV604 DNA was digested for 45 minutes with 1 unit (lane 3), 0.5 units (lane 4), 0.25 units (lane 5), 0.125 units (lane 6), 0.06 units (lane 7) or 0.03 units (lane 8) of *Eco*RI. For comparison undigested (lane 2) and fully digested (lane 10) pPCV604 DNA were run. These contain open circle plasmid (upper band, lane 2), supercoiled plasmid (lower band, lane 2) and the fully digested 8 kb band (upper band, lane 10). Lanes 1 and 9 are *Hin*dIII plus *Hin*dIII/*Eco*RI digested λ DNA. Numbers on the left of the figure indicate band sizes in kb.

plasmid running below and relaxed circular plasmid running above. The 8 kb fully digested band could not be avoided. The DNA was then phenol extracted before digestion with *Hin*dIII followed by calf alkaline phosphatase treatment and phenol purification. This complex of DNA fragments was expected to contain four fragments; 9 kb *Hin*dIII-*Hin*dIII, 8 kb *Eco*RI-*Hin*dIII, 8 kb *Eco*RI-*Eco*RI and the desired 9 kb *Eco*RI-*Hin*dIII (refer Figures 1 and 7).

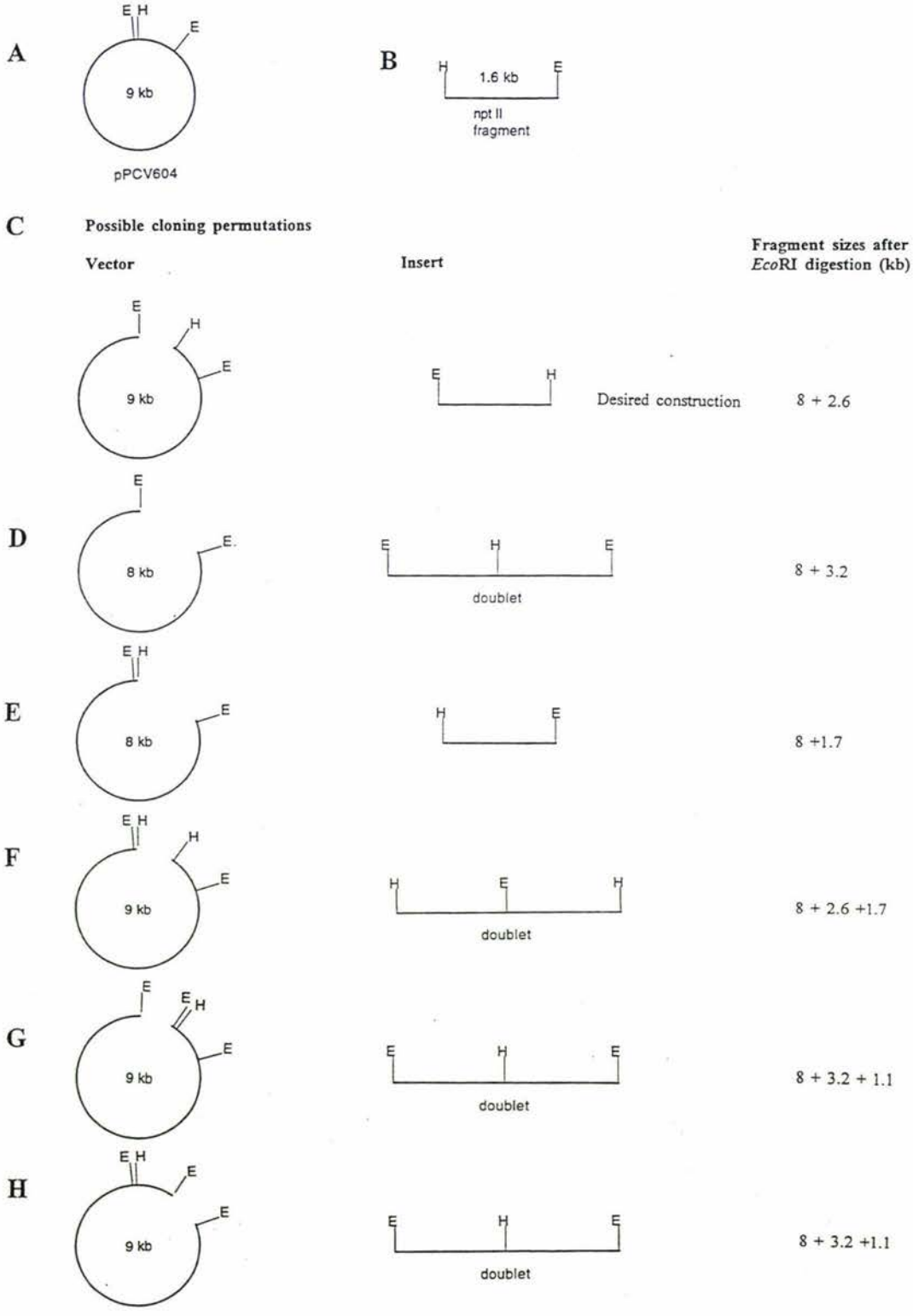
The pPCV604 fragments were ligated (Section 2.3.9) with the 1.6 kb *Eco*RI/*Hin*dIII fragment from pNNPT, transformed into S17-1 (Section 2.3.10) and colonies selected on LB 50 mg l^{-1} ampicillin. All the transformants screened, with *Hin*dIII and *Eco*RI single and double digests, were identical to pPCV604. It was postulated that uncut plasmid remained after digestion with *Hin*dIII, perhaps due to inhibition of the enzyme by carry over of phenol or other inhibitors from the gel extraction process (Section 2.3.13). Another possibility was that the alkaline phosphatase activity was inhibited, so that digested molecules were free to religate, though a greater range of ligation products would have been expected.

In the absence of a blue/white selection process using β-galactosidase α-complementation to screen against the parental plasmid, it was necessary to ensure that both *Hin*dIII and alkaline phosphatase activity were optimal. Because of the small fragments involved (100 base pairs and 1 kb) and the mix of fragments from the *Eco*RI partial digest, it was not possible to check the efficiency of *Hin*dIII digestion if used second. Hence *Hin*dIII digestion was conducted first for one hour and a sample removed to ensure complete digestion by observation after gel electrophoresis. Then *Eco*RI was added at the predetermined optimum concentration for partial digestion (0.125 units in a 45 minute incubation), making the assumption that predigestion with *Hin*dIII would not significantly alter the rate of *Eco*RI digestion. Calf alkaline phosphatase (Section 2.3.8) was added for the final 20 minutes of the digest before phenol purification (Section 2.3.5).

The pPCV604 fragments and the 1.6 kb pNNPT fragment were ligated, transformed into S17-1 and colonies selected on LB 50 mg l^{-1} ampicillin. Transformations with pPCV604 fragments unligated, pPCV604 fragments self ligated, and pNNPT fragment self ligated were conducted as controls. A tenth of the volume of the ligation

Figure 7. Potential outcome of cloning during construction of pGT.

Plasmid pPCV604 (**A**) was *Hin*dIII and partial *Eco*RI digested and ligated to a 1.6 kb fragment from pNNPT (**B**). Potential ligations of variously cut vector and insert combinations and their restriction digestion products with *Eco*RI are shown (**C** to **H**). The desired ligation is displayed in **C**.


A
E H
E
9 kb
pPCV604
B
H
1.6 kb
E
npt II
fragment
C
Possible cloning permutations
Vector
Insert
Fragment sizes after EcoRI digestion (kb)
E
H
E
9 kb
E
H
Desired construction
8 + 2.6
D
E
E
8 kb
E
H
E
doublet
8 + 3.2
E H
E
8 kb
H
E
8 +1.7
F
E H
H
E
9 kb
H
E
H
doublet
8 + 2.6 +1.7
G
E
E H
E
9 kb
E
H
E
doublet
8 + 3.2 + 1.1
H
E H
E
E
9 kb
E
H
E
doublet
8 + 3.2 +1.1

reactions was sampled before and after ligase treatment and observed following gel electrophoresis (Figure 8). There was no observable change in the banding of the pPCV604 fragments pre-ligation compared to post-ligation (Figure 8, lanes 2 and 3) suggesting that alkaline phosphatase treatment was preventing self ligation. Ligation of the 1.6 kb pNNPT fragment alone resulted in the 1.6 kb band (Figure 8, lane 6, pre-ligation) being replaced with a faint 3.2 kb band (Figure 8, lane 7, post-ligation), presumably from duplex formation. The ligation designed to create pGT (Figure 8, lane 4, preligation; lane 5, post-ligation) also resulted in disappearance of the 1.6 kb band, but the addition of higher molecular weight bands running at approximately 9 and 15 kb. Together, this suggested that a significant proportion of ligation products were due to ligation occurring between pNNPT and pPCV604 fragments, with possible closed circular (running higher) and linear 9.5 kb ligation products responsible for the new bands (Figure 8, lane 5).

The ligations were transformed into S17-1 and colonies selected on 50 mg l^{-1} ampicillin. The results of this transformation are summarised in Table 4. No transformants resulted from self ligation of the pNNPT fragment, suggesting there was insignificant carry over of pUC8 derived vector. But there were as many, or more transformants from unligated and self ligated pPCV604 fragments (Table 4) as in the cloning ligation. This suggested that, again, undigested pPCV604 had remained and dominated the transformant colonies.

Twenty colonies from the cloning ligation were picked and plasmid DNA screened with *Eco*RI digests (Figure 9). A range of the simplest possible ligation products are listed in Figure 7. Four isolates had bands identical to pPCV604 (8 and 1.1 kb, Figure 9, lanes 3, 9, 17 and 24), six had the desired restriction pattern (8 and 2.6 kb, Figure 9, lanes 2, 6, 7, 13, 19 and 23) while the remainder had either 8 and 1.7 kb bands (Figure 9, lanes 5, 8, 10, 11, 12, 18, 20, 21 and 22), or 9, 8 and 1.7 kb bands (Figure 9, lane 4). Four lanes (1, 14, 16 and 25) carried *Eco*RI digests of the parent vector pPCV604 for comparison.

The six potential isolates initially found with 8 and 2.6 kb *Eco*RI bands were checked with *Hin*dIII-*Eco*RI double digests for 8, 1.6 and 1.0 kb bands and one selected and named pGT. This was further restriction mapped (Figure 10).

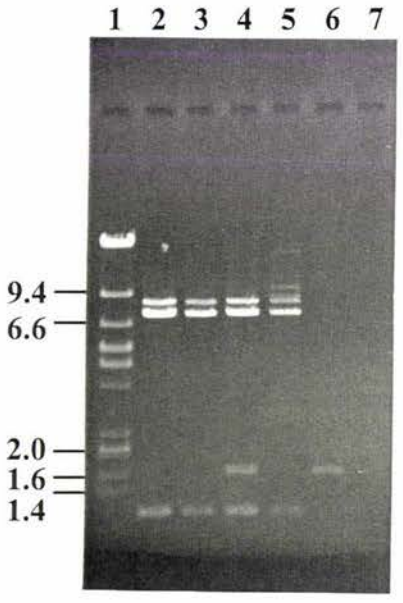


Figure 8. Ligation reactions and controls for construction of pGT.

Lane 1: *Hin*dIII plus *Hin*dIII/*Eco*RI digested λ DNA size standards (Section 2.3.12)

Lane 2: Preligation of *Hin*dIII, partial *Eco*RI digested pPCV604.

Lane 3: Postligation of *Hin*dIII, partial *Eco*RI digested pPCV604.

Lane 4: Preligation of mixed *Hin*dIII, partial *Eco*RI digested pPCV604 DNA and 1.6 kb *npt*II fragment from pNNPT.

Lane 5: Postligation of mixed *Hin*dIII, partial *Eco*RI digested pPCV604 DNA and 1.6 kb *npt*II fragment from pNNPT.

Lane 6: Preligation of 1.6 kb *npt*II fragment from pNNPT

Lane 7: Postligation of 1.6 kb *npt*II fragment from pNNPT

Numbers on the left of the figure indicate band sizes in kb.

Table 4. Transformation of S17-1 after DNA ligation to create pGT.

DNA	Ligase treated	Fraction of transformation mix plated onto LB 50 mg l^{-1} ampicillin		
		1/20	1/4	1/2
pPCV604 fragments[1]	no	0	nd[3]	17
pPCV604 fragments[1]	yes	0	3	9
pPCV604[1] plus pNNPT[2] fragments	yes	0	7	14

[1] Plasmid pPCV604 DNA was linearised with *Hin*dIII, partially digested with *Eco*RI and alkaline phosphatase treated.

[2] A 1.6 kb *Eco*RI/*Hin*dIII fragment cut from pNNPT.

[3] nd: not determined.

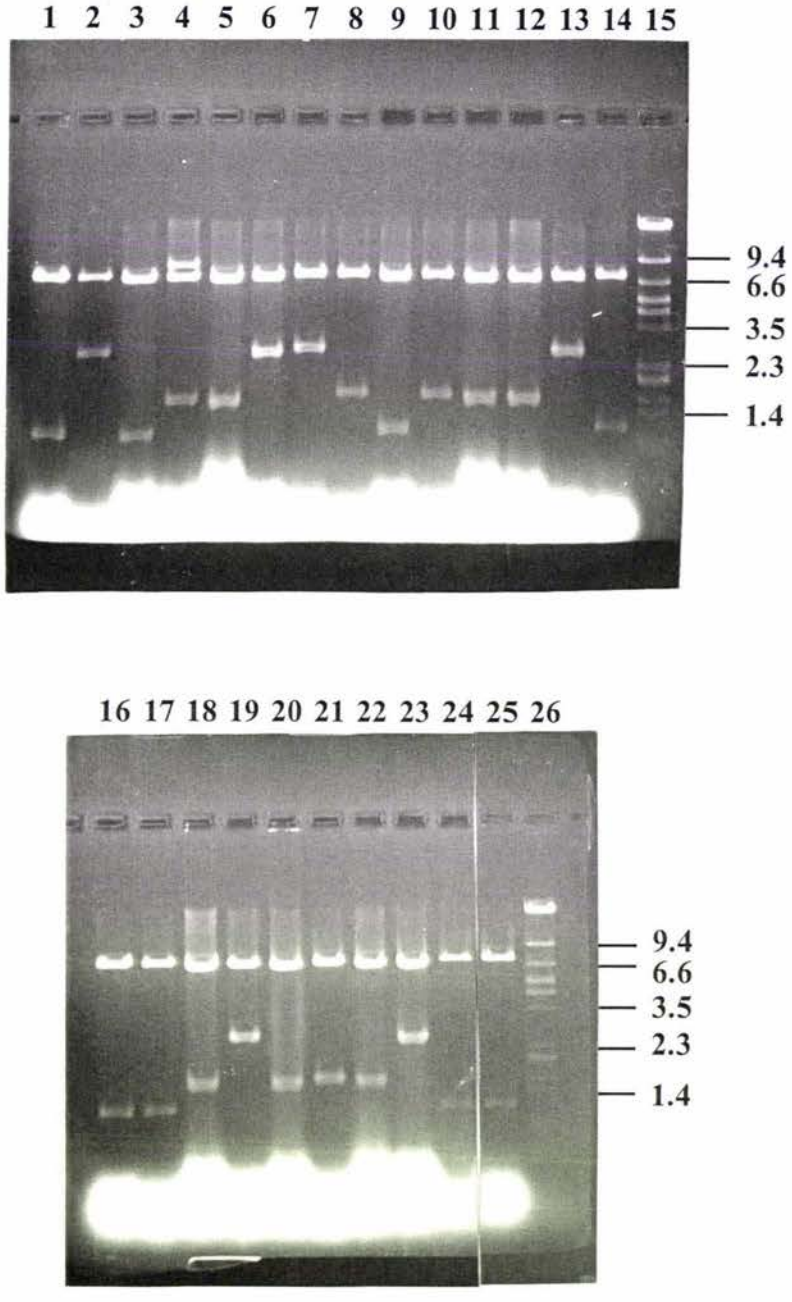


Figure 9. Screening of putative pGT clones with *Eco*RI.

Lanes 1, 14, 16 and 25: *Eco*RI digest of pPCV604 DNA.

Lanes 2 to 13 and 17 to 24: *Eco*RI digests of DNA from transformation isolates.

Lanes 15 and 26: *Hin*dIII plus *Hin*dIII/*Eco*RI digested λ DNA size standards (Section 2.3.12).

Numbers on the right of the figure indicate band sizes in kb.

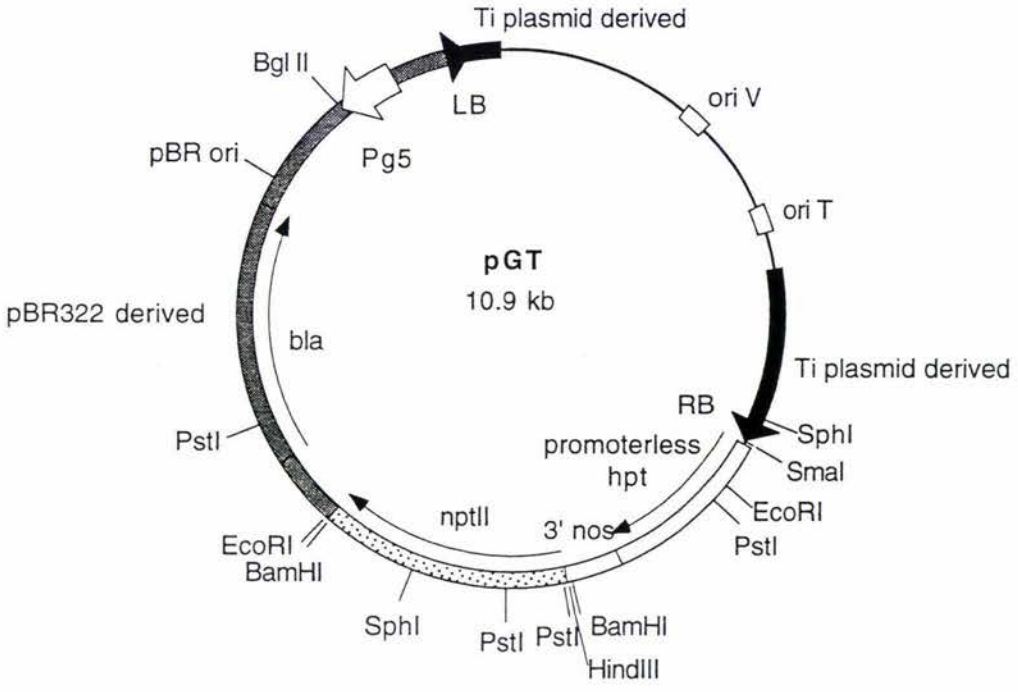


Figure 10. Physical map of pGT.

Abbreviations: LB and RB, left and right borders of T-DNA; pBR ori, replication origin from pBR322; *hpt*, coding region of hygromycin phosphotransferase gene; *npt*II, chimaeric Pnos−*npt*II−3′*nos* gene; *bla*, β-lactamase gene; Pg5, promoter T-DNA gene 5.

The new plasmid pGT (Figure 10) is a promoter tagging vector in its own right and contains a new selectable marker, the chimaeric *nos-npt*II, independent of the promoterless reporter gene *hpt* used for both tagging and selection of transformed plants in the parent vector pPCV604. However to obtain a more sensitive promoterless reporter gene system the promoterless *hpt* of pGT was now replaced with the promoterless β-glucuronidase (*gus*) gene to create pGTG.

3.1.2 Construction of pGTG

The *gus* gene including *ocs* 3′ processing signals but excluding the promoter sequence was excised from pKIWI101a (Janssen and Gardner, 1989). This fragment appeared as the largest of four bands after digestion with *Xho*I and *Bam*HI. This DNA was extracted from a gel and ligated into *Sal*I/*Bam*HI cut pUC8 (Appendix 8.1.2). Transformants were selected on 50 mg l^{-1} ampicillin and XGal and one isolate was picked and the plasmid named pUCK1 (Figure 11). The *gus* gene was excised from pUCK1 with *Sal*I (site derived from pKIWI101a sequence) and *Hin*dIII, gel extracted, and recloned into *Sal*I/*Hin*dIII cut pUC8, making pUCK2 (Figure 12), which was selected on ampicillin as for pUCK1 and screened by restriction digestion checking for the loss of the *Pst*I site.

The plasmid pUCK2 could have been created in one cloning step directly by cloning the *Xho*I-*Sal*I *gus* gene fragment from pKIWI101a directly into pUC8 cut with *Sal*I and then screening for the correct orientation. However, the two step cloning did not require orientation checking and pUCK1 was desired for other purposes. On the other hand, the next cloning step, taking a pUCK2 *Sma*I-*Hin*dIII fragment and cloning into pGT, could have similarly been done from pUCK1. But this would have added additional nucleotides between the right border and the *gus* start codon. These would have included a spurious ATG codon within the *Sph*I sequence of pUCK1 that was likely to lead to attenuated expression of the reporter gene (Bevan, 1984).

The *gus* gene was then cut out of pUCK2 (Figure 12) with *Sma*I and *Hin*dIII, separated by gel electrophoresis and purified with glassmilk (Geneclean, Section 2.3.13) and ligated into similarly digested and purified pGT (Figure 10). The latter

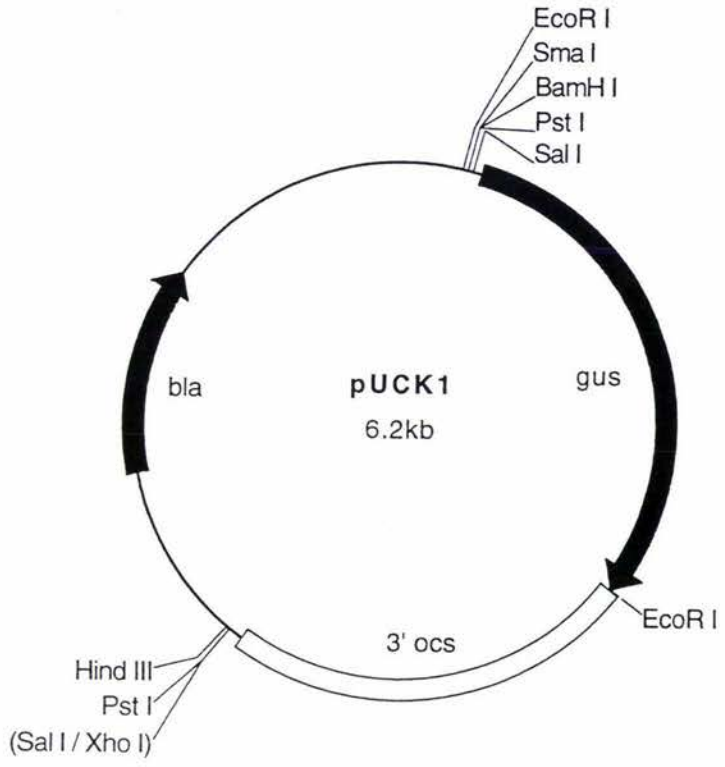


Figure 11. Physical map of pUCK1.

Plasmid pUCK1 contains the 3.5 kb *Bam*HI/*Xho*I fragment from pKIWI101a (Janssen and Gardner, 1989) containing promoterless β-glucuronidase (*gus*) gene with octopine synthase terminator sequence (3′ *ocs*) ligated into *Bam*HI/*Sal*I cut pUC8. The *Sal*I/*Xho*I ligation removed the ability of these sites to be digested.

Abbreviation: *bla*, β-lactamase;

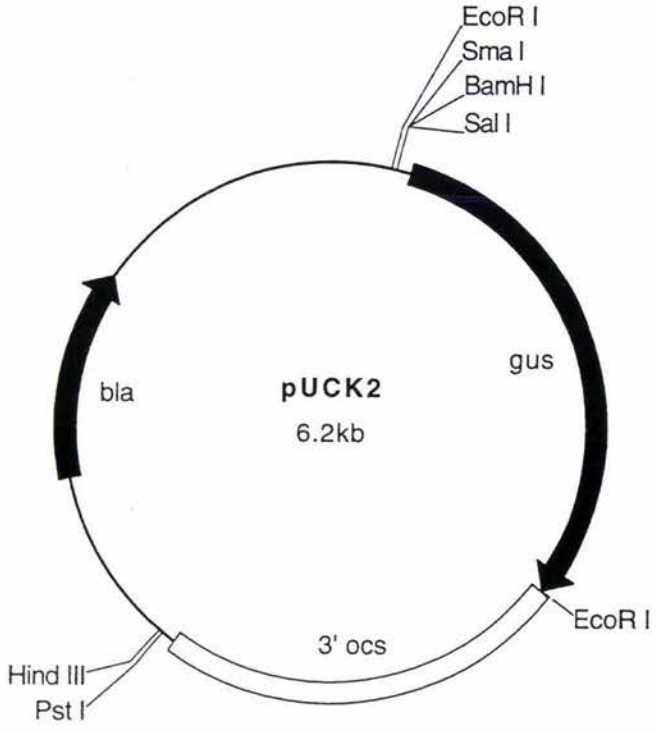


Figure 12. Physical map of pUCK2.

Plasmid pUCK2 contains the 3.5 kb *Sal*I/*Hin*dIII fragment from pUCK1 with the promoterless β-glucuronidase (*gus*) gene and octopine synthase terminator sequence (3′ *ocs*) cloned into *Sal*I/*Hin*dIII cut pUC8.

Abbreviation: *bla*, β-lactamase.

digest removed the promoterless *hpt* gene including 3′ *nos* processing signals from pGT. The resulting plasmid was restriction mapped (Table 5) and named pGTG (Figure 13).

To confirm the construction of pGTG (Figure 13), two fragments were subcloned into identically cut vectors with pUC18 and pUC19 polylinkers and partially sequenced from one end with universal primer (Section 2.3.15). These fragments were the 1.9 kb *Sph*I-*Eco*RI fragment spanning the T-DNA right border to the 3′ end of the *gus* coding sequence and a 3.2 kb *Bam*HI-*Eco*RI fragment from the 3′ end of the *gus* coding sequence to the 3′ end of the chimaeric P*nos*-*npt*II-3′*nos* cassette.

The sequence obtained from the *Sph*I site across the right border and into the *gus* coding sequence (Figure 14) contained two regions with exact matches to Genbank database sequences (Section 2.7.6), namely to the T-DNA right border region of pTiC58 (Genbank accession number J01819) and also to the *gus* coding region (Genbank accession number M14641). Several important features which have been preserved from the parent vector pPCV604 are highlighted on the sequence (Figure 14). The overdrive core sequence (Peralta *et al.*, 1986) is located just downstream of the *Sph*I site. The 25 base pair right border element finishes 62 bases upstream of the *gus* start codon. This is exactly the same distance between the right border and the promoterless reporter gene start codon as in pPCV604 between the right border and the promoterless *hpt* gene. As in pPCV604 there are stop codons in all three reading frames between the right border sequence and the *gus* start codon so that any insertion within the coding sequence of a plant gene cannot result in a translational fusion with *gus*, but only a transcriptional fusion with two separate translated reading frames.

The three other sequences obtained from the ends of the two fragments subcloned from pGTG matched the expected Genbank database sequence. Sequence determined from the *Eco*RI end of the 1.9 kb fragment matched *gus* coding sequence (Genbank accession number M14641). Sequence at the *Bam*HI end of the 3.2 kb fragment matched 3′ *nos* sequence (Genbank accession number J01541) and at the *Eco*RI end of the 3.2 kb fragment matched 3′ *ocs* sequence (Genbank accession number J01820). Hence these sequences supported the assumption that pGTG had been correctly assembled.

Table 5. Restriction mapping of pGTG.

Restriction digest	Fragment size (kb)
*Eco*RI	9.7, 3.2
*Bam*HI	7.8, 5.1
*Pst*I	10.6, 1.9, 0.4
*Eco*RI/*Bam*HI	7.8, 2.2, 1.9
*Eco*RI/*Hin*dIII	9.7, 1.6, 1.6
*Hin*dIII/*Bam*HI	7.8, 3.5, 1.6
*Sph*I	8.2, 4.7
*Sph*I/*Eco*RI	7.6, 2.6, 2.1, 0.6
*Bgl*II	12.9
*Eco*RI/*Bgl*II	6.7, 3.2, 3.0

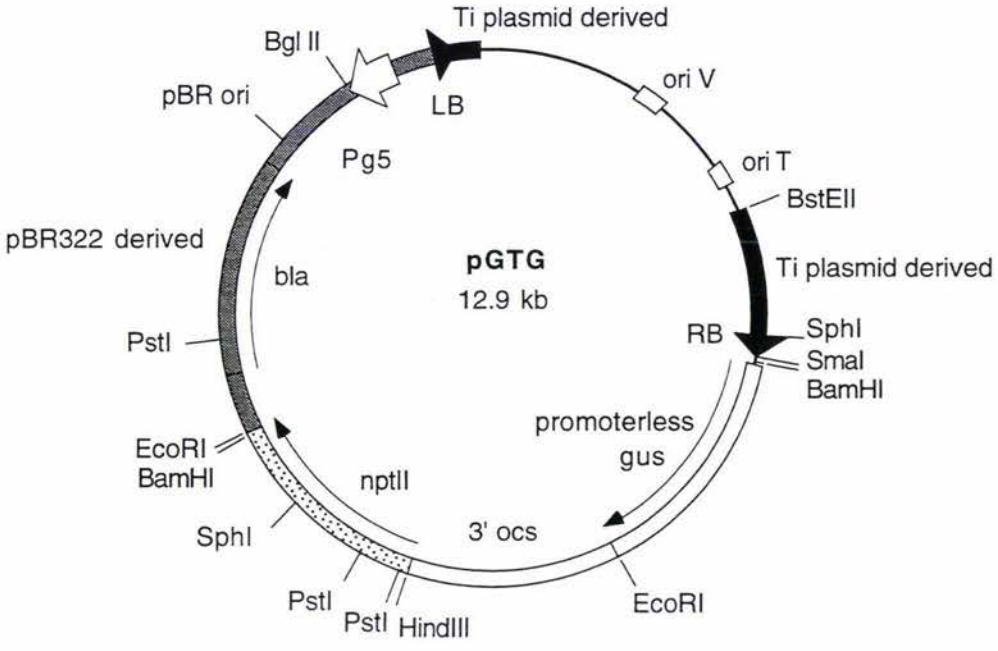


Figure 13. Physical map of pGTG.

Abbreviations: LB and RB, left and right borders of T-DNA; pBR ori, replication origin from pBR322; *gus*, coding region of β-glucuronidase gene; 3′ *ocs*, octopine synthase termination sequence; *npt*II, chimaeric Pnos−*npt*II−3′nos gene; *bla*, β-lactamase gene; Pg5, promoter T-DNA gene 5.

```
     overdrive
 1   cacataca aa tggacgaacg gataaacctt ttcacgccct tttaaatatc
SphI

                                                  right border
51   cgattattct aataaacgct cttttctctt aggtttaccc gccaatatat

                  STOP CODONS
101  cctgtcaaac acTGATAGtt TAAaccgaag gcgggaaacg acaatcTGAt

                                * *
151  cccggggatc cgtcgaccat ggtccgtcct gtagaaaccc caacccgtga
     SmaI BamHI SalI     start codon

201  aatcaaaaaa ctcgacggcc tgtgggcatt cagtctggat cgcgaaaact

251  gtggaattga tcagcgttgg tgggaaagcg cgttacaaga aagccgggca

301  attgctgtgc caggcagttt taacgatcag ttc
```

Figure 14. DNA sequence determined across the T-DNA right border of pGTG.

An *Sph*I-*Eco*RI fragment subcloned from pGTG was sequenced from the *Sph*I end. Bases 1 to 138 had perfect homology to the right border of pTiC58 (Genbank accession number J01819). Bases 169 to the end had perfect homology to the expected *gus* sequence. The two bases immediately after the start codon labelled with asterisks differ from the Genbank database *gus* sequence (accession number M14641) due to subsequent manipulations (Jefferson, 1987) designed to improve the translational efficiency of *gus* in a eukaryote system. Restriction endonuclease sites are underlined. The core overdrive sequence is coloured green. The right border is coloured magenta. The *gus* start codon is coloured blue. Stop codons between the right border and the *gus* start codon are capitalised and coloured red.

The promoter tagging vector, pGTG, now had the required promoterless *gus* reporter gene, but it retained features of pPCV604 that were not desired. The first was a host range limited to strains providing pRK2 replication functions in *trans* of which only *Agrobacterium tumefaciens* strain GV3101 was available and this strain had a low tobacco transformation efficiency (Section 3.2). The second was the inclusion of the T_L gene 5 promoter from pTiAch5 (Koncz and Schell, 1986) just inside the left border facing internally. It seemed appropriate in a promoter tagging vector to remove unnecessary plant regulatory elements, especially one known to act in a tissue-specific manner (Koncz and Schell, 1986). This would prevent any possibility of enhancer activity or transcriptional interference (Ingelbrecht *et al.*, 1991) acting on *gus* expression, which could disrupt the reporter gene from accurately reflecting the expression pattern of any tagged promoter.

3.1.3 Construction of pBin19-GTG

It was determined that the right border and immediately external sequence of pPCV604 (Figure 1; derived from pTiC58; Koncz and Schell, 1986; Koncz, 1989) and pBin19 (Appendix 8.1.3; derived from pTiT37; Bevan, 1984) were likely to be identical (Zambryski *et al.*, 1980). This included two restriction sites external to the right border, *Sph*I and *Bst*EII. It appeared possible to join the replicon features of pBin19 to the T-DNA of pGTG (Figure 13). Hence a fragment beginning at one of these sites, just outside the T-DNA right border, and ending at the *Bam*HI site within the polylinker just inside the left border of pBin19 could be ligated to a right border and internal T-DNA fragment of pGTG extending to the *Bgl*II site, stopping before both the left border and the T_L gene 5 promoter.

It was determined that both *Sph*I and *Bst*EII cut twice within pGTG, the second site within the T-DNA. However, *Sph*I fortuitously cut preferentially at the right border site. Hence a partial *Sph*I and *Bgl*II digested and purified (Section 2.3.13) 8.3 kb T-DNA fragment of pGTG was ligated to the *Sph*I-*Bam*HI replicon fragment of pBin19. The new plasmid pBin19-GTG (Figure 15) was confirmed by restriction mapping (Table 6).

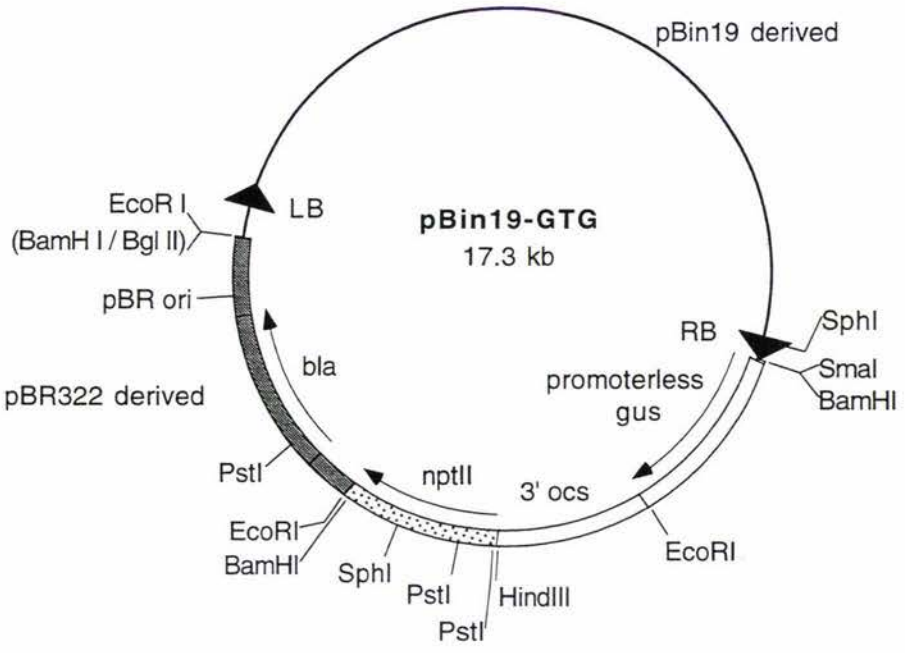


Figure 15. Physical map of pBin19-GTG.

Abbreviations: LB and RB, left and right borders of T-DNA; pBR ori; replication origin from pBR322; *gus*, coding region of β-glucuronidase gene; 3′ *ocs*, octopine synthase 3′ processing signals; *npt*II, chimaeric Pnos−*npt*II−3′*nos* gene; *bla*, β-lactamase gene.

Table 6. Restriction mapping of pBin19-GTG.

Restriction digest	Fragment size (kb)
*Hin*dIII	17.3
*Eco*RI	11.1, 3.2, 3.0
*Pst*I	15, 1.9, 0.4
*Eco*RI/*Hin*dIII	11.1, 3.0, 1.6, 1.6
*Bam*HI	12.2, 5.1
*Bam*HI/*Pst*I	11.5, 3.5, 1.2, 0.7, 0.4
*Eco*RI/*Bam*HI	9.2, 3.2, 3.0, 1.9
*Hin*dIII/*Bam*HI	12.2, 3.5, 1.6

3.2 Comparison of tobacco transformation with pGTG and pBin19-GTG

The two promoter tagging vectors, pGTG and pBin19-GTG are very similar, except that pGTG is limited to host strains providing pRK2 replication functions in *trans* of which only *Agrobacterium tumefaciens* strain GV3101 was available. Plasmid pBin19-GTG is a standard broad host range vector able to replicate in any *Agrobacterium tumefaciens* strain. The tobacco transformation efficiencies of pGTG in GV3101 and of pBin19-GTG in LBA4404 (in common use within the laboratory) were compared.

Both *gus* promoter tagging vectors were crossed into *Agrobacterium* strains; pGTG into GV3101 and pBin19-GTG into LBA4404. *Nicotiana tabacum* (DSIR line KKD) was transformed (Section 2.4.5) using the protocol of Horsch *et al.* (1985).

First, it was confirmed that equivalent optical density of cultures of the two strains equated to equivalent numbers of colony forming units by dilution plating. Subsequently, the optical density of the strains was equalised when resuspending the cells in $MgSO_4$ immediately prior to tobacco inoculation. The leaf pieces were plated on NicI for a two day co-cultivation period and then onto NicII 300 mg l^{-1} kanamycin with 100 or 500 mg l^{-1} cefotaxime. The number of calli clumps and shoots were counted after 22 days (Table 7). Less than half the leaf pieces had callus growth at this time following GV3101/pGTG inoculation, while after LBA4404/pBin19-GTG inoculation over 90% of leaf pieces had callus growth. There was no apparent difference between GV3101/pGTG on either 100 or 500 mg l^{-1} cefotaxime. However, there were more than double the number of calli regions on the lower cefotaxime concentration with LBA4404/pBin19-GTG as well as 12 strong shoots against none with any other treatment.

After another four weeks growth, the leaf pieces inoculated with GV3101/pGTG had become overgrown with bacteria, but there was no sign of bacteria on any other treatment. There was callus growth on every leaf piece in every treatment except for two leaf pieces inoculated with GV3101/pGTG on 500 mg l^{-1} cefotaxime. However the regeneration was clearly stronger after LBA4404/pBin19-GTG inoculation with 22 large shoots (greater than 5 mm high with distinct leaves) on 100 mg l^{-1} cefotaxime

Table 7. Comparison of tobacco regeneration[1] with different *A. tumefaciens* strain and binary vector combinations.

A. tumefaciens strain and binary plasmid	Cefotaxime concentration (mg l^{-1})	Total number of leaf pieces inoculated	Average number of calli regions per leaf piece	Number of leaf pieces with stated number of regions of callus					Number of strong shoots (greater than 5 mm high)
				0	1	2-5	6-10	>10	
GV3101/pGTG	100	20	0.6	13	5	2	0	0	0
GV3101/pGTG	500	20	0.5	12	6	2	0	0	0
LBA4404/ pBin19-GTG	100	18	11.6	0	0	2	10	6	12
LBA4404/ pBin19-GTG	500	18	4.6	3	1	8	5	1	0

[1] Leaf pieces were inoculated with equivalent densities of each strain of bacteria and co-cultivated for two days on NicI before placing on NicII 300 mg l^{-1} kanamycin and either 100 or 500 mg l^{-1} cefotaxime. Regions of regeneration on each tobacco leaf piece were counted after 3 weeks.

and 16 on 500 mg l^{-1} cefotaxime, while only two strong shoots had emerged on each of the GV3101/pGTG inoculations.

Figure 16 is a photograph of a similar experiment three weeks after inoculation with leaf pieces plated on NicII 300 mg l^{-1} kanamycin, 100 mg l^{-1} cefotaxime. Numerous regions of calli can be seen on the LBA4404/pBin19-GTG inoculated leaf pieces (Figure 16 **A**), and very few on the GV3101/pGTG inoculated leaves (Figure 16 **B**). Bacterial growth can be seen affecting two of these leaf pieces (Figure 16 **B**).

Leaf pieces damaged by forceps during inoculation were more susceptible to GV3101 overgrowth and with this strain there was markedly more bacterial growth after co-cultivation. On the other hand, this strain was much less tolerant of antibiotic selection and therefore it was grown on 25 mg l^{-1} kanamycin and 50 mg l^{-1} carbenicillin compared to 100 mg l^{-1} of both these antibiotics for LBA4404.

Further LBA4404/pBin19-GTG inoculated leaf pieces were grown on 100 mg l^{-1} cefotaxime while GV3101/pGTG transformations were plated on 500 mg l^{-1} cefotaxime. Transgenic plants generated with both these constructs were then screened for GUS activity.

3.3 Fluorometric screening for GUS activity

Transgenic tobacco plants growing in tissue culture in individual pottles were analysed at approximately similar stages of growth, about two weeks after cut shoots rooted on NicIII media containing kanamycin and cefotaxime. A fully expanded leaf and several lengths of root including root tips and lateral roots were fluorometrically assayed for GUS activity on microtitre plates. An example of this type of assay is illustrated in Figure 17.

From 89 plants tested, 33 plants (37%) were found to have fluorescence indicative of GUS activity. Four of these had expression only in the leaf material, 16 only in roots and the remaining 13 expressed GUS in both leaf and roots. Eleven of these plants had detectable fluorescence after one hour. These were from one plant with

Figure 16. Comparison of *Nicotiana* transformation by GV3101 with pGTG and LBA4404 with pBin19-GTG.

Plates of *Nicotiana tabacum* leaf pieces on NicI Km 300 mg l^{-1} Cf 100 mg l^{-1} were photographed three weeks after inoculation with *Agrobacterium*.

A LBA4404/pBin19-GTG inoculated.

B GV3101/pGTG inoculated. Bacterial growth can be seen on the upper left leaf piece.

A

B

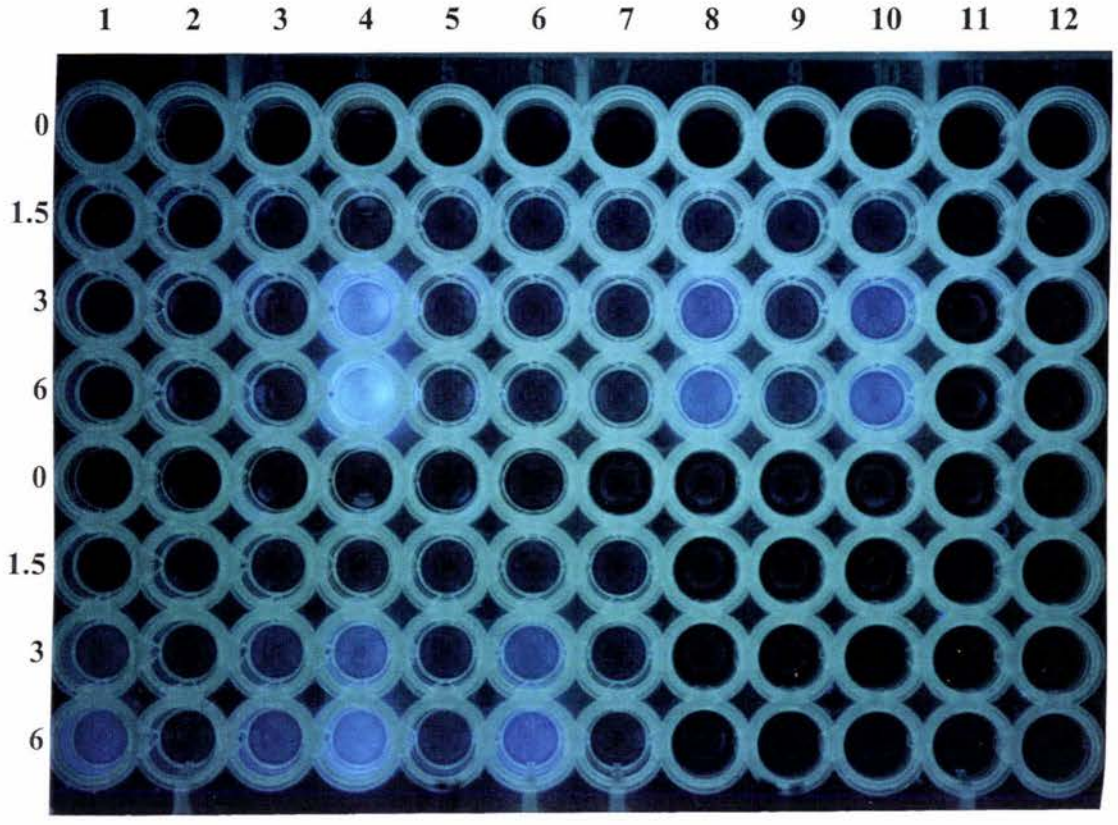


Figure 17. Microtitre plate screening for fluorometric GUS activity.

Plant extracts (25 µl) were mixed with 25 µl of assay buffer. Reactions were stopped at four time points (0, 1.5, 3 and 6 hours; numbers on the left of the figure) with 50 µl of 0.5 M Na_2CO_3. The microtitre plate was photographed under ultraviolet illumination. Protocol details are in Sections 2.5.3 and 2.5.4. From Kerr *et al.* (1991).
Column 4, upper four rows: positive control from P_{35S}-*gus*-3′ *ocs* transgenic plant.
Columns 5 to 7, upper four rows: E4 seedling 1; stem extract (column 5), root extract (column 6) and leaf extract (column 7).
Columns 8 to 10, upper four rows: E4 seedling 2; stem extract (column 8), root extract (column 9) and leaf extract (column 10).
Columns 1 to 3, lower four rows: E4 seedling 3; stem extract (column 1), root extract (column 2) and leaf extract (column 3).
Columns 4 to 6, lower four rows: E4 seedling 4; stem extract (column 4), root extract (column 5) and leaf extract (column 6).
Remaining wells are empty.

expression only in the leaf, three plants with expression only in the roots and six plants with both leaf and root expression, though in three of these plants the strong expression was confined to the leaf material.

Controls without the GUS substrate, MUG, were routinely run. It was noted that older roots showed significant levels of MUG-independent fluorescence. This activity increased with time during the assay indicating an enzymatic process rather than pre-existing fluorescent material. Hence care was taken to distinguish this background level from any increased fluorescence with MUG added indicative of GUS activity.

The two plants (E1 and 118) with the strongest levels of root specific GUS expression were selected for further study.

3.4 GUS activity in plants E1 and 118

3.4.1 Fluorometric assay

The GUS activity in root and leaf samples from plants E1 and 118 grown in tissue culture was determined fluorometrically (Section 2.5.5).

Plant E1 had GUS activity of 1330 (± 60) pmol MU per mg protein per minute (± to maximum range of values) in root tissue and 3 (± 3) pmol MU per mg protein per minute in leaf tissue. Plant 118 had GUS activity of 1200 (± 70) pmol MU per mg protein per minute in root tissue and 1 (± 1) pmol MU per mg protein per minute in leaf tissue.

The GUS activity in the roots of both transgenic plants was much higher than the background average of 6 (± 5) pmol MU per mg protein per minute in untransformed root extracts, though the leaf measurements were not significantly different from the average of 3 (± 3) pmol MU per mg protein per minute in untransformed leaf extracts.

This indicated, as the microtitre plate assay also had, that plant E1 had higher GUS expression than plant 118 in root material and that there was no expression apparent

above background fluorescence in the leaf material of either plant.

In contrast, histochemical staining (Section 3.4.2) revealed much stronger *gus* expression in plant 118 relative to plant E1. As well, the level of expression was variable, though becoming less intense as time progressed. After a few months there was little GUS activity detectable in tissue culture plants, though it remained strong in glasshouse grown plants. As glasshouse root material could not be assayed without contamination with soil micro-organisms which had significant GUS activity, new tissue culture lines of both plants were regenerated on Nic media from discs from surface sterilised leaves.

Fluorometric GUS testing of these lines revealed only three out of fourteen 118 clones were GUS positive in root tissue (average 230 ± 130 pmol MU per mg protein per minute), while all seven E1 clones were GUS positive in root material (average 230 ± 160 pmol MU per mg protein per minute). The presence or absence of 118 GUS activity was independent of whether the plant material had been regenerated in the presence or absence of antibiotics (kanamycin and cefotaxime). Leaf material was always GUS negative.

3.4.2 Histochemical staining

Leaf and root samples were cut from tissue culture grown plants E1 and 118 and stained for GUS activity using XGluc (Section 2.5.1). After destaining the material was observed under a dissecting microscope. Leaf and stem material from both plants was always found to be negative. However root material produced distinctive and reproducible patterns of GUS activity, indicative that the promoterless *gus* reporter gene was now being transcribed, presumably in the patterns of endogenous plant genes that had been subject to T-DNA insertions.

Plant 118 had strong GUS activity in lateral root tips, the sites of lateral root attachment to the main root (Figure 18 **A**) and also in and around root primordia (Figure 18 **B**). The activity in root tips was intense throughout the root cap and the meristematic region, fading in colour moving back 3 or 4 mm along the root

Figure 18. GUS activity in plants 118 and E1.

A and **B** Roots from plant 118.

C Root tip from plant 118.

D Older root from plant 118.

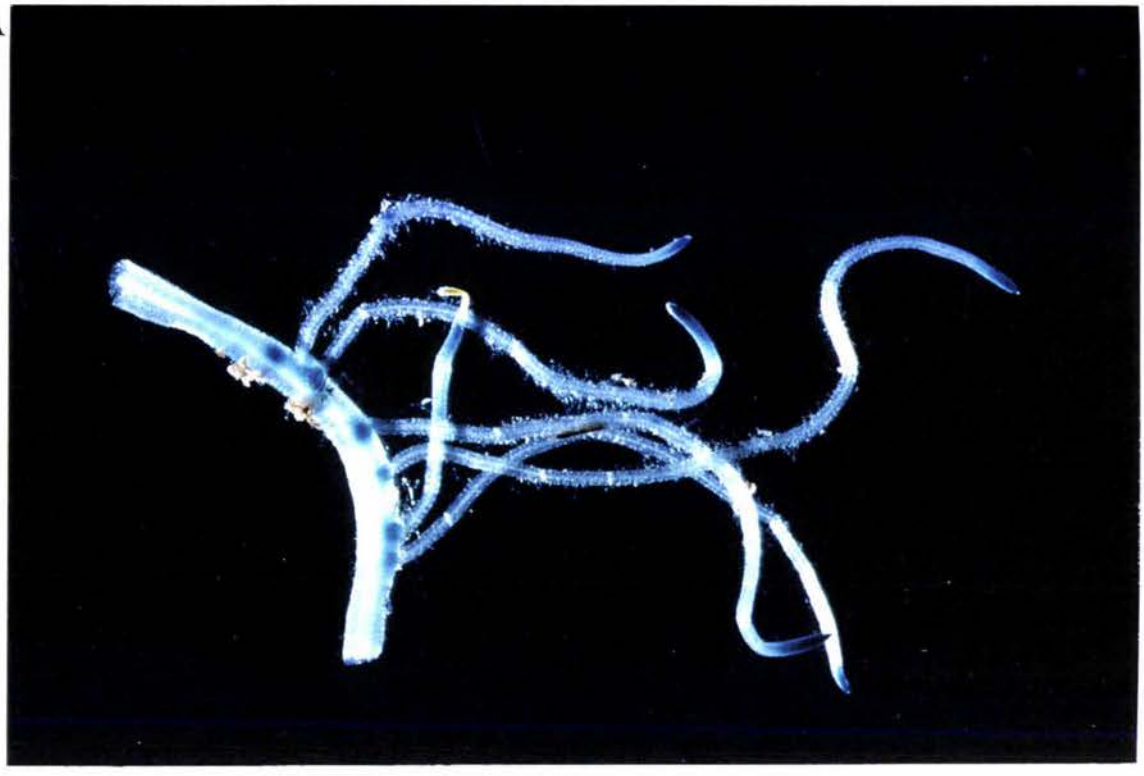
E Comparison of E1 root tip (entering photograph from bottom left) and 118 root tip (centre, entering photograph from top right).

F Root tip of older lateral root than in **E** from plant E1.

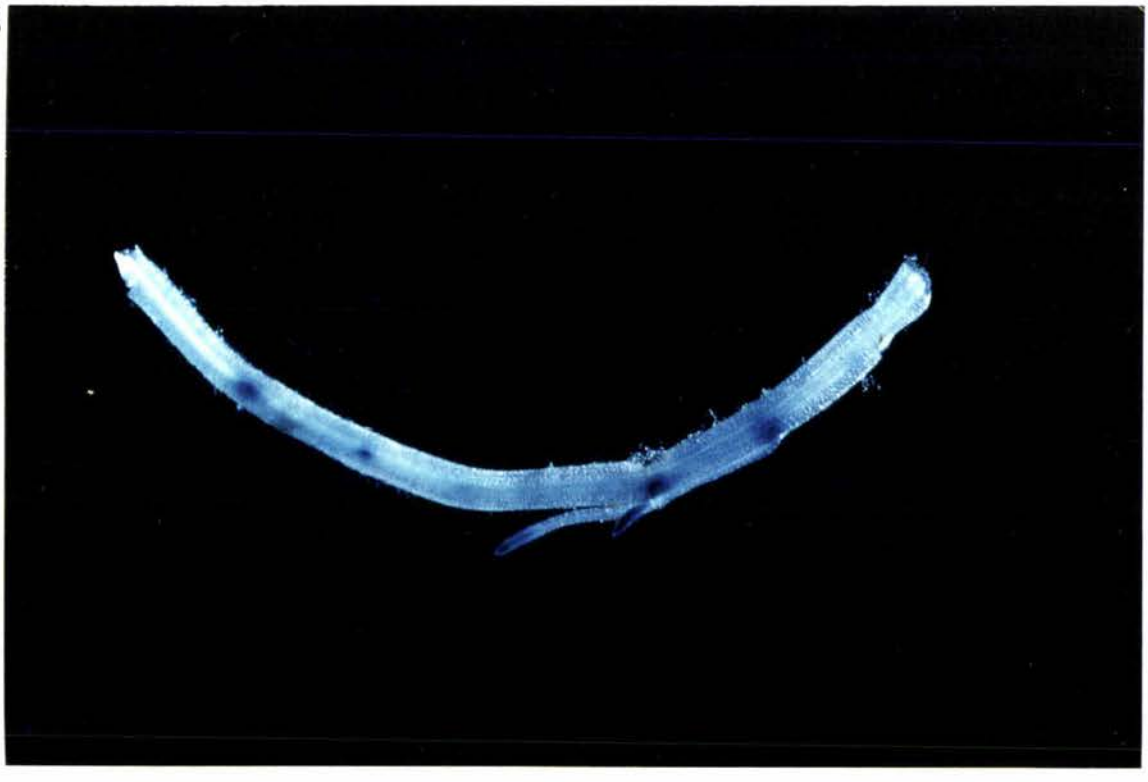
G and **H** Roots from plant E1.

Photographs **A**,**C**, **D**, **F** and **G** from Kerr *et al.* (1991).

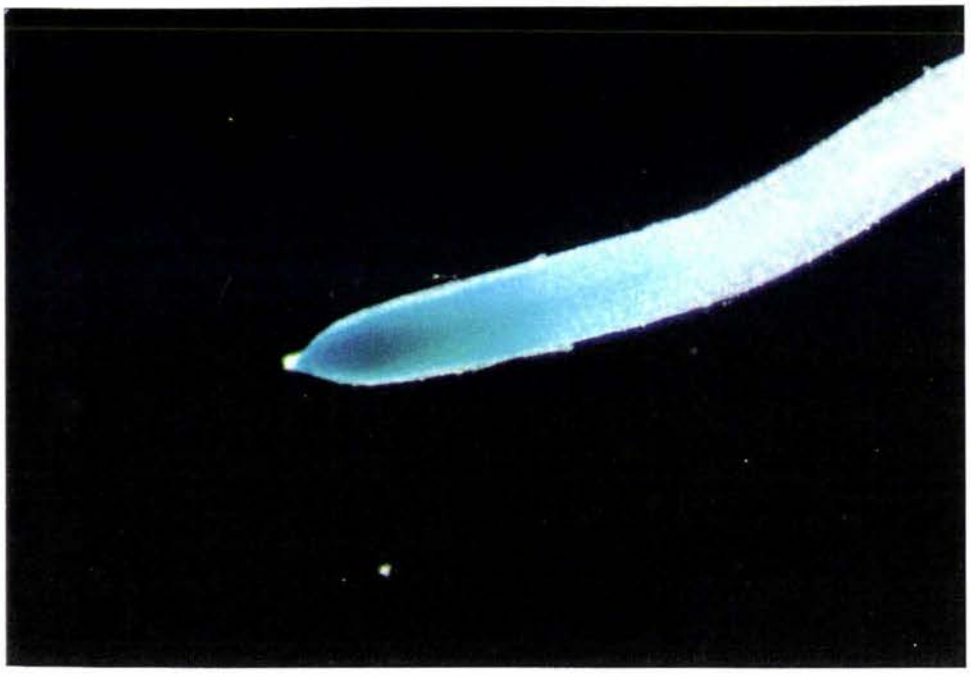
A



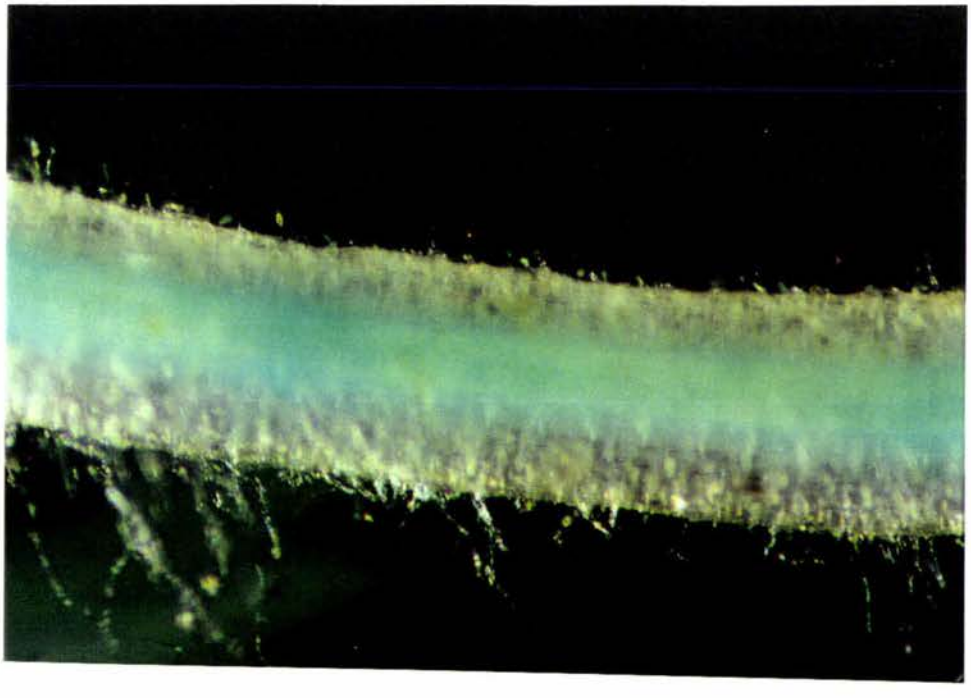
B



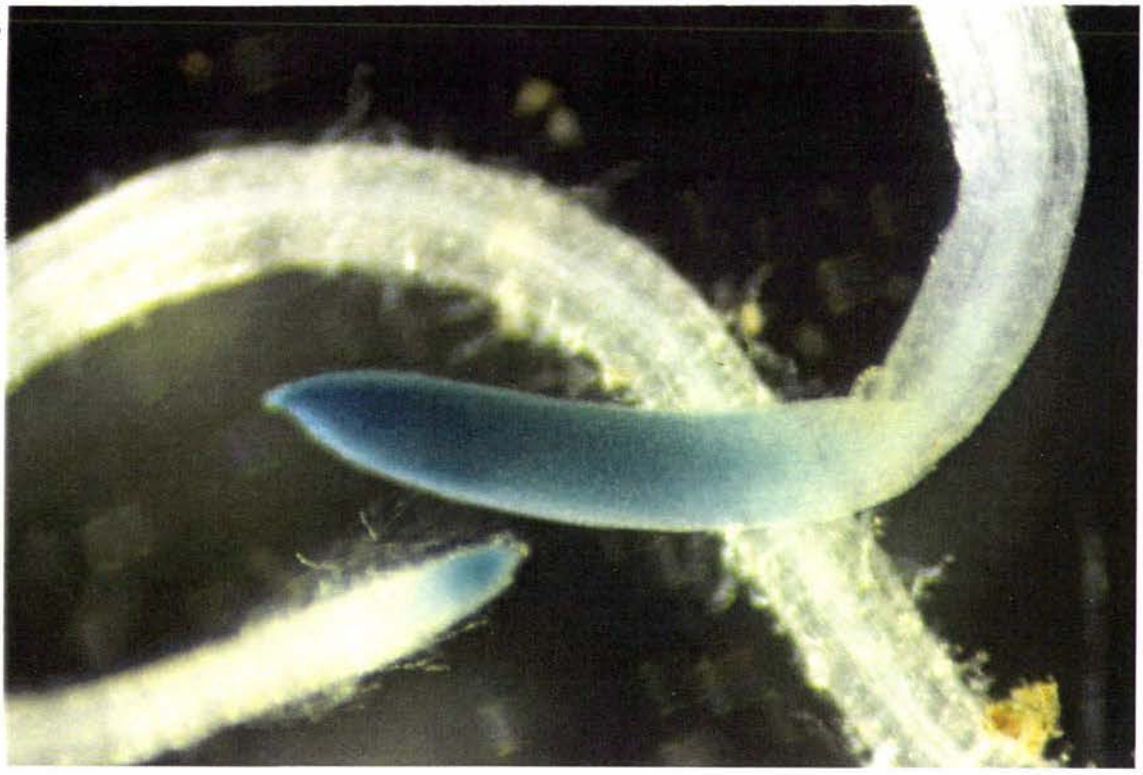
C



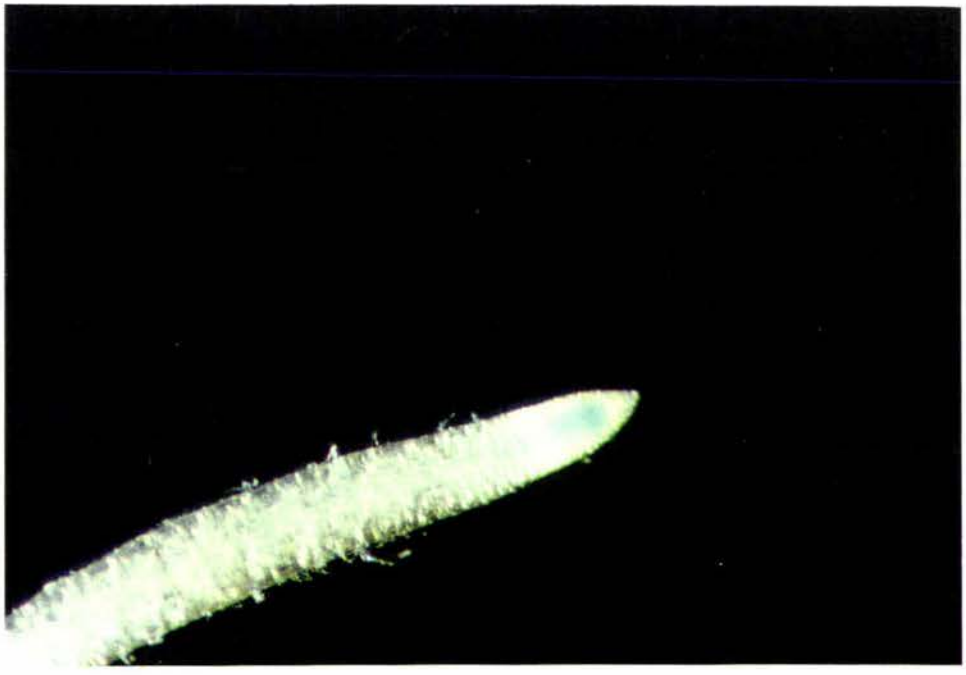
D



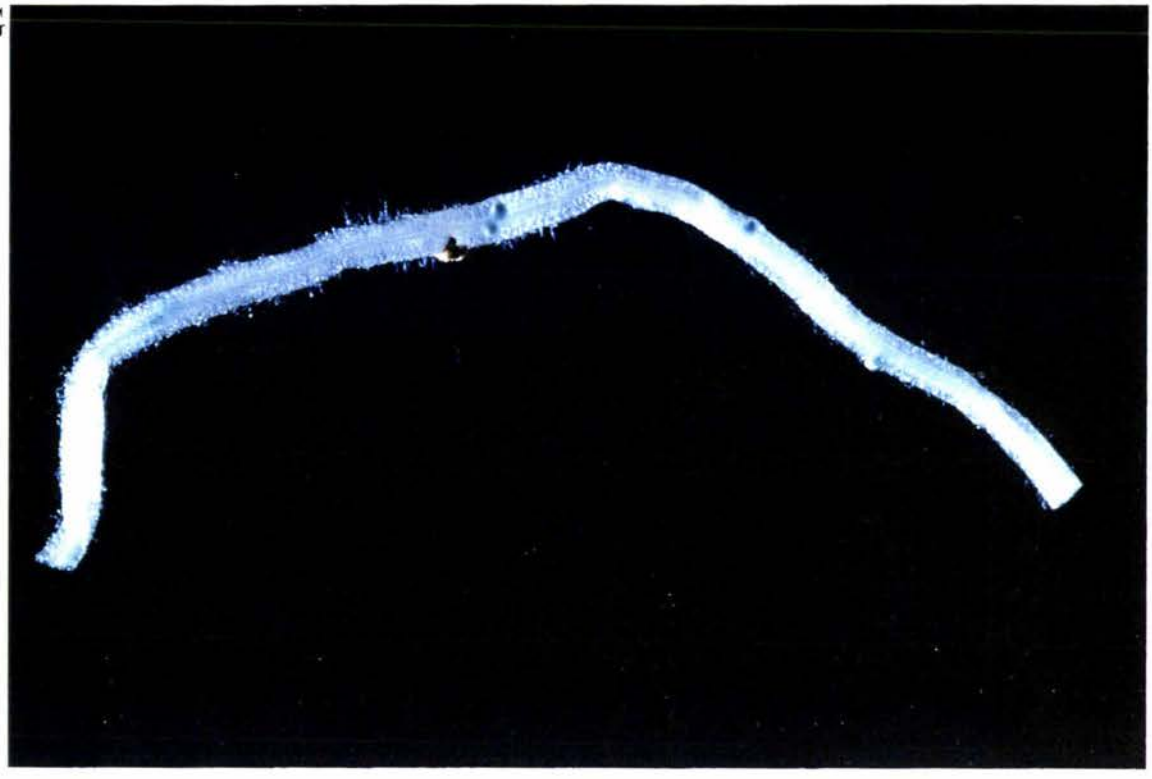
E



F



G



H

(Figure 18 **C**). Staining was weak and patchy in root vascular tissue but occasionally an older root was seen with moderate GUS activity strictly confined to the stele (Figure 18 **D**).

Comparing the GUS activity of root tips of plant E1 to plant 118 (Figure 18 **E**), it was noted that the activity in a young root tip of plant E1 was weaker and much more tightly confined to the root apex. In older lateral roots the intensity of XGluc staining was much weaker and seen only in the apical meristem and procambial strand immediately adjacent (Figure 18 **F**). There was no staining at all in the root tips of older (longer) lateral roots (data not shown), however staining was apparent in lateral root primordia (Figure 18 **G** and **H**), though unlike plant 118 (compare Figure 18 **B** with **G**) this GUS activity did not penetrate into cortical cells around the root primordia. Nor did the GUS activity remain at the branching point (data not shown).

As detailed above (Section 3.4.1) GUS activity decreased in tissue culture grown material. Because of the variable nature of the expression, possibly influenced by the tissue culture growth conditions (for example, older or more crowded roots may have less expression), it was decided to analysis GUS activity in seedlings following self-fertilisation of both plants. This would allow the analysis of multiple replicates at identical developmental stages and also allow the expression pattern to be traced through the full plant life cycle.

3.5 Analysis of progeny

3.5.1 GUS expression

Seeds from self-pollination of the previously chosen plants E1 and 118 with root specific GUS activity were surface sterilised and grown on 1/2 MS and water agar. The seedlings growing on both media were histochemically tested for GUS activity. Unexpectedly, no staining was seen.

In order to check that silencing in progeny was not a feature of this tobacco variety (DSIR line KKD) nor of the promoter tagging vector system, seedlings from another

strongly GUS positive plant, E4, were tested and some were found to be positive (Figure 19). The staining intensity could readily be classified into three groups, strong, weak and none (Figure 19) with a ratio of 3:6:2 respectively. The pattern of staining (Figure 20) was identical in all the GUS positive seedlings. The strongest staining was seen in the stem and axillary buds with weaker staining throughout the leaves. Root staining was confined to root tips, lateral root initials and around the branch points.

3.5.2 Azacytidine treatment

Seeds obtained after self-pollination of plants E1 and 118 were germinated on water agar containing 0, 5 or 30 μM 5-azacytidine. After three weeks, 20 seedlings from each plate were histochemically tested for GUS expression. The only positive staining was seen at the highest 5-azacytidine concentration with plant 118 seedlings (Figure 21). Seven were positive, 12 negative and one unscorable (root broken off when sampled). Among the positive plants the staining intensity was variable and was never as great as in the parent plant.

At six weeks more seedlings were tested from each treatment and both E1 and 118 batches. All were found to be GUS negative.

A second trial was set up, with 0, 30, 50 and 100 μM azacytidine in water agar. Ten 118 seedlings were tested from each plate after 3 weeks. On 0 μM 5-azacytidine all ten were GUS negative, on 30 μM 5-azacytidine seven plants were positive and on both 50 and 100 μM 5-azacytidine six plants were GUS positive. The higher concentrations of azacytidine made no difference to the intensity of the staining, though the seedlings had stunted growth.

Overall, 26 seedlings from plant 118 were GUS positive out of 59 tested on 30, 50 or 100 μM 5-azacytidine (44%).

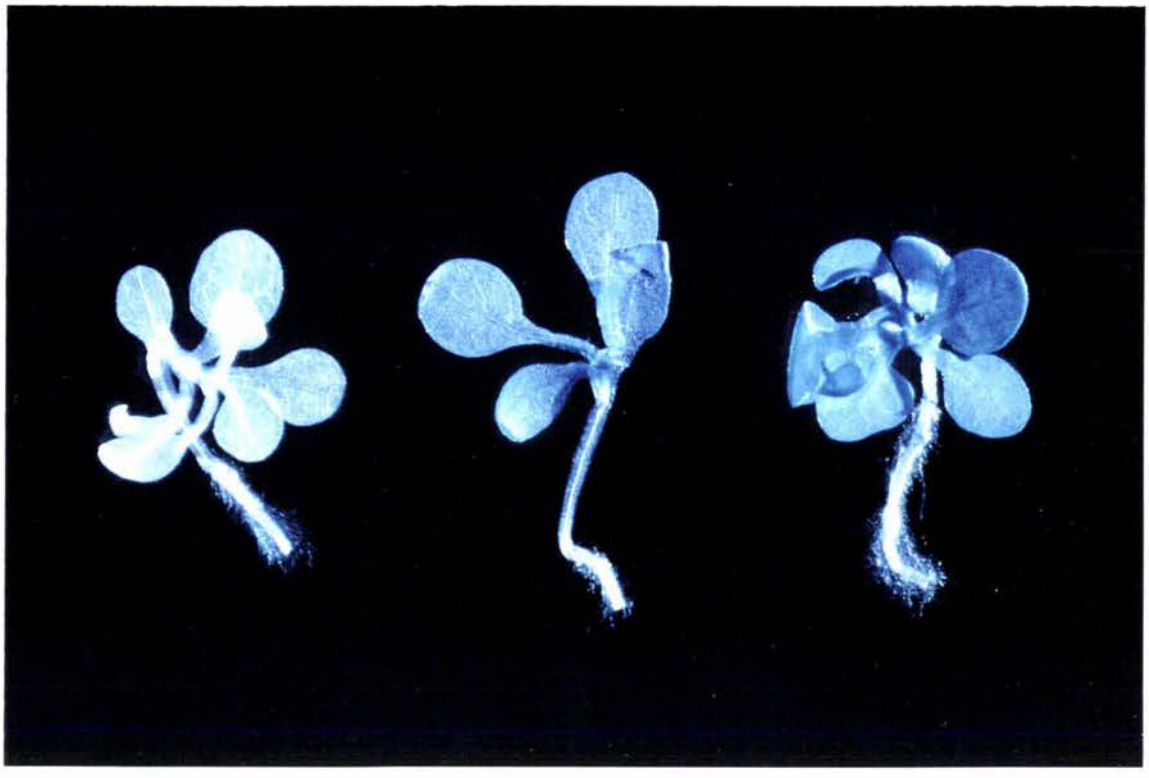


Figure 19. GUS activity in seedlings of plant E4.

After histochemical staining with XGluc, seedlings of plant E4 could be grouped into three categories: No staining (seedlings on left), moderate staining (centre seedling) or strong staining (seedling on right). From Kerr *et al.* (1991).

A

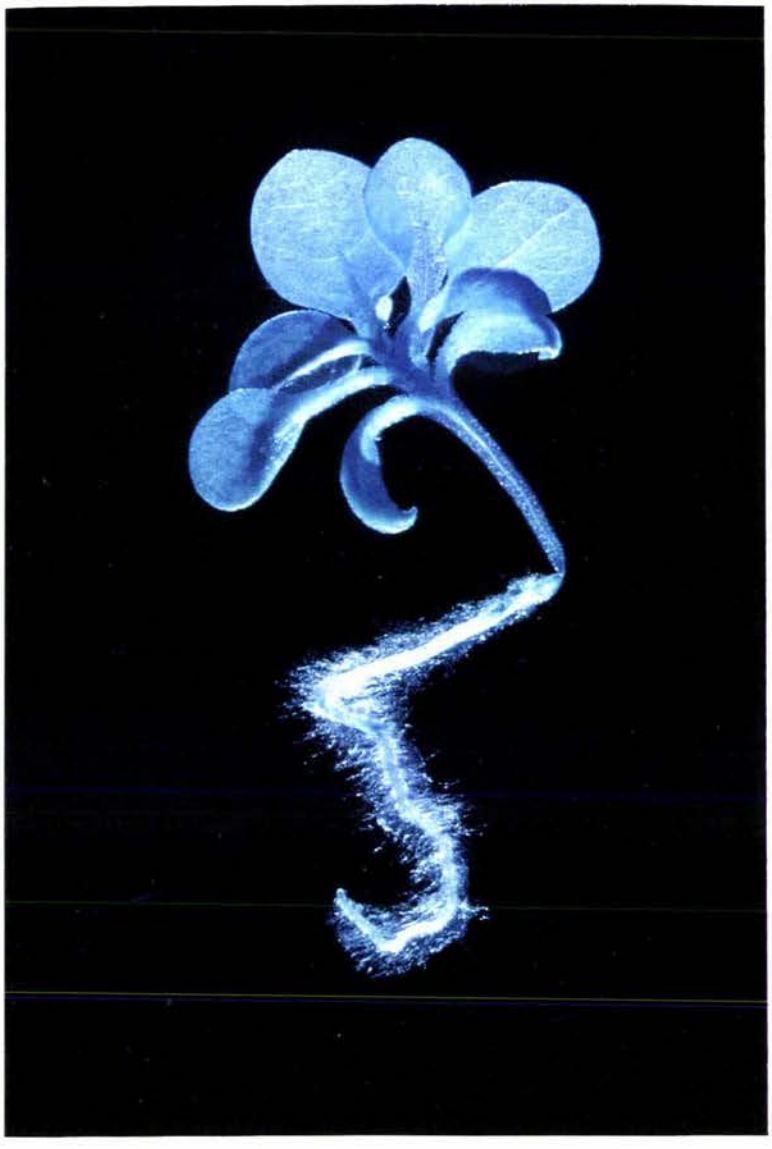


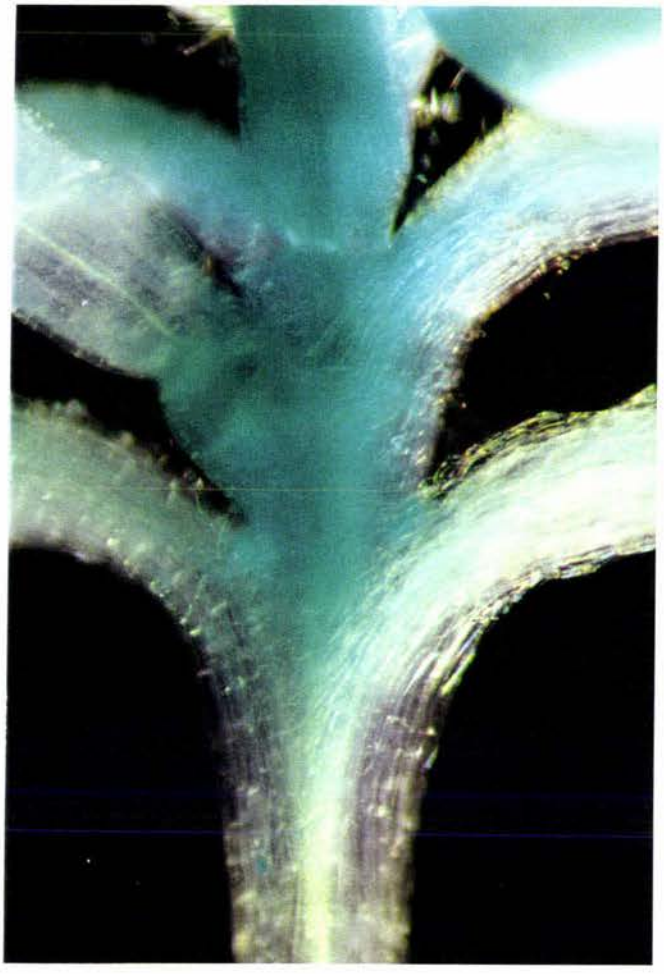
Figure 20. Localisation of GUS activity in a seedling of plant E4.

From Kerr *et al.* (1991).

A: Seedling.

B: Stem of same plant.

B



A

Figure 21. GUS activity of 118 seedlings grown on 5-azacytidine.

Surface sterilised seedlings of plant 118 were grown on water agar containing 30 μM 5-azacytidine. After 3 weeks seedlings were histochemically tested for GUS activity and photographed unbleached.

A Selection of all the GUS positive and some negative seedlings.

B One GUS positive seedling.

B

3.5.3 Kanamycin resistance

The kanamycin resistance of seedlings from self-pollinated transgenic plants were tested in order to estimate the number of independent T-DNA copies.

Kanamycin resistance of seedlings was tested on a variety of media. On water agar with kanamycin all seedlings quickly bleached whilst on MS media the seedlings were very slow to bleach. On 1/2 and 1/4 MS, with kanamycin 200 mg l^{-1}, control seedlings (Figure 22 **A**) bleached, their epicotyl failed to grow and their root only penetrated 1 to 2 mm into the media. These three characteristics were used to score kanamycin sensitive seedlings (for example Figure 22 **B**) from plants E1, E4 and 118 (Table 8). Assuming Mendelian segregation, this indicates that E1 has one kanamycin resistance gene, 118 has three independently segregating kanamycin resistance genes and E4 has two.

Scoring was not always clear cut. While control seedlings were always distinctly affected by kanamycin, transgenic seed appeared to give a gradation of sensitivity. Occasionally only cotyledons would emerge from the seed coat, or the root would not penetrate the media but instead point into the air. These seedlings were excluded from assessment.

Experiments were not always scorable. This was due to non-uniform growth both among seedlings on one plate and between plates. The main reason for this was variable germination of seed, especially those of a previous season's seed batch. Slower germinating seedlings could be difficult to distinguish from kanamycin sensitive seedlings. The water content of individual plates of media also influenced the germination rate and the growth of seedlings. Different genotypes of tobacco also responded differently. Wisconsin-38 plants remained greener for longer on kanamycin than DSIR line KKD plants, though the distinguishing characteristics of sensitive seedlings held true.

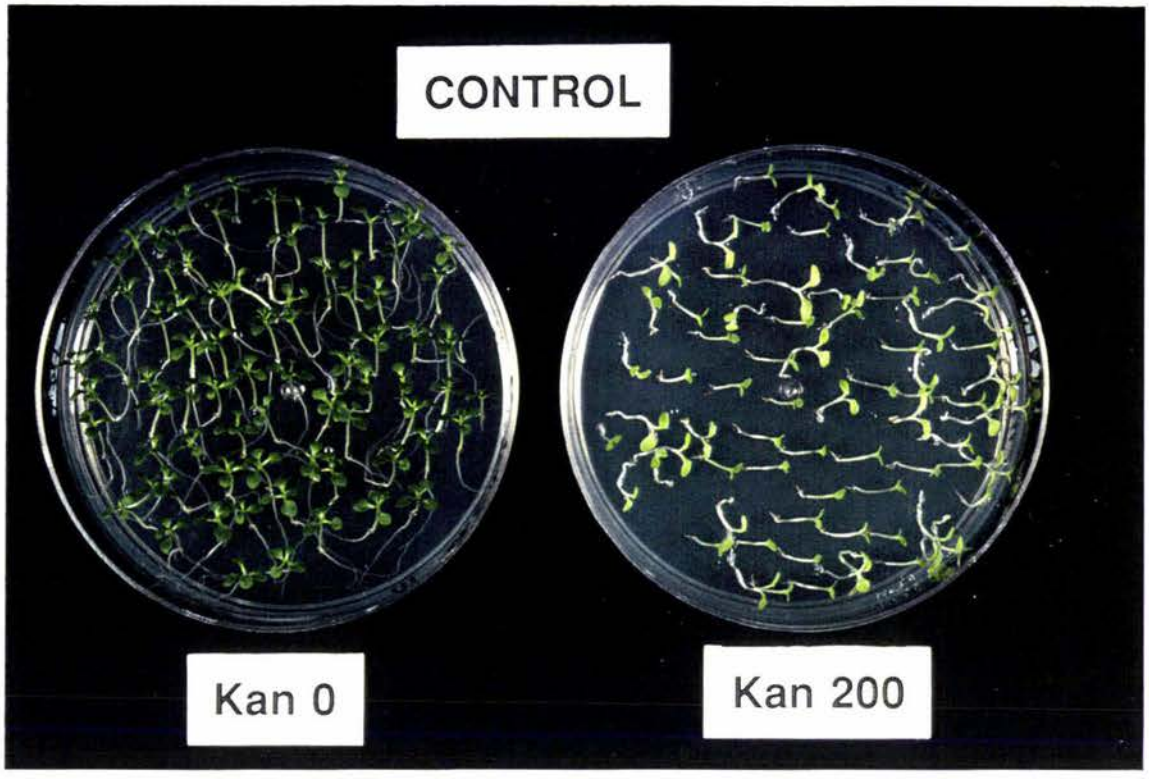
Figure 22. Screening for kanamycin resistance.

Surface sterilised seed were plated onto water agar containing 0 or 200 mg l^{-1} kanamycin. The plants were photographed after 1 month.

A Untransformed *Nicotiana tabacum*.

B Plant E1.

A



B

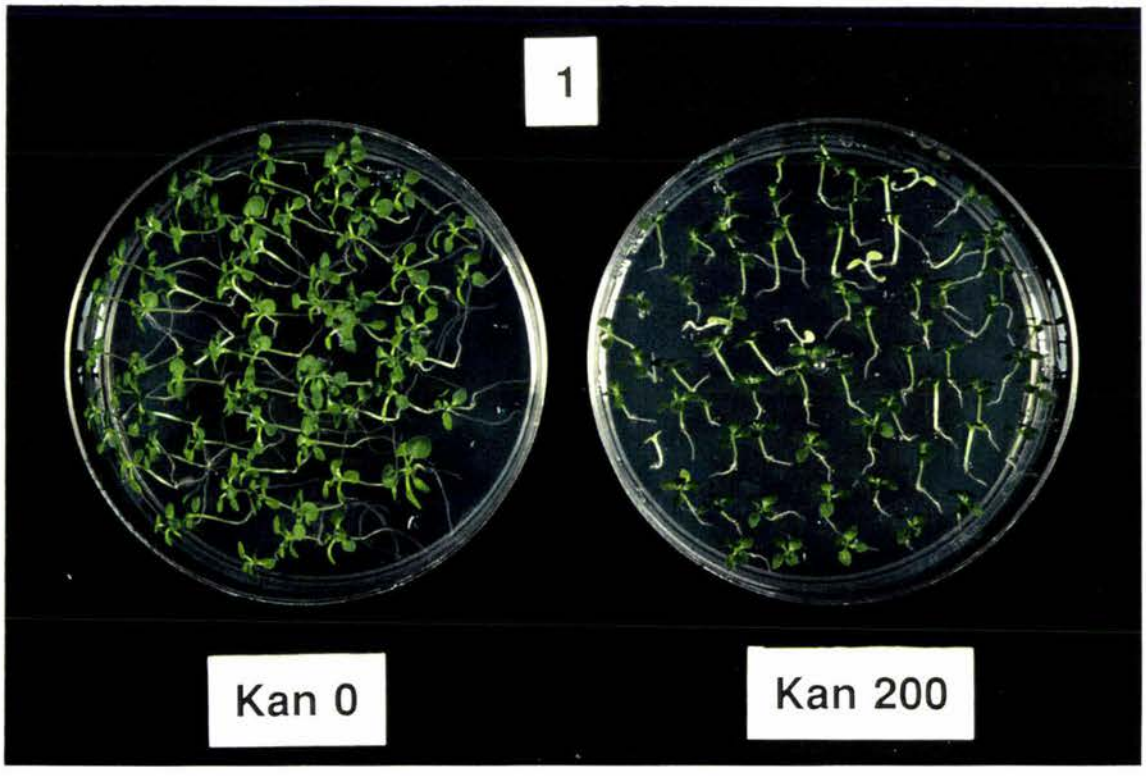


Table 8. Kanamycin sensitivity of transgenic progeny.

Plant	Number of sensitive seedlings	Total number of seedlings screened	Best Mendelian ratio of sensitive seedlings to total number of seedlings	Chi-squared probability
E1	20	130	1:3	P = 0.8
118	3	245	1:63	P = 0.975
E4	3	50	1:15	P = 0.99

3.6 Southern hybridisation

Southern hybridisation was used to assess T-DNA copy number in transgenic plants as well as providing a guide for the cloning of plant genomic sequence flanking T-DNA insertions.

Genomic DNA (20 μg) from plants E1 and 118 was digested with *Eco*RI, ethanol precipitated, separated by electrophoresis, Southern blotted and hybridised against a probe of *gus* prepared by PCR with universal and reverse primers from a subclone of pUCK1 containing only the *gus* coding sequence.

A full length, and single, copy of the T-DNA of either pBin19-GTG (plant 118) or pGTG (plant E1) integrated into tobacco genomic DNA should generate a restriction fragment that hybridises to the *gus* probe which is greater than 1.9 kb after *Eco*RI digestion.

Autoradiography (Figure 23) revealed bands of 10, 8.5, 3.8, 3.1, 3.0, 2.8 and 2.5 kb in the digested 118 DNA (lane 3) and bands of 6.3, 3.8, and 2.6 kb in the E1 digested DNA (lane 2).

With this probe, all these fragments should have one end generated from the internal T-DNA *Eco*RI site just downstream of the *gus* coding sequence. Their other *Eco*RI site should be generated from upstream of the right border, hence within flanking plant genomic sequence or within other T-DNA sequence assuming a rearrangement or multiple insertion. Hence these bands provide an estimate of T-DNA insertions in both plants, seven in plant 118 and three in plant E1.

The intensity of the 3.8 kb band, seen in the 118 *Eco*RI digest (Figure 23, lane 3), is suggestive that it is present in multiple copies. These could have been created by several inverse repeats of the T-DNA around the right border (ie doublets of the 1.9 kb *gus* fragment) or from direct repeats of the T-DNA with a small deletion near the left border. This band migrates more slowly than the intensely *npt*II-hybridising 3.5 kb band seen in these two plants (data not presented) and others transformed with these plasmids (for example refer Section 4.4 Figure 37 **C**, lanes 3 to 7) which is an

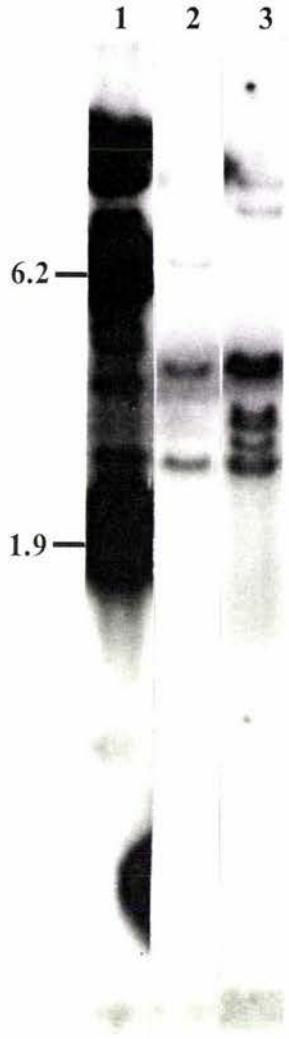


Figure 23. Autoradiograph of Southern hybridisation of genomic DNA from plants E1 and 118.

DNA isolated from plants E1 (lane 2) and 118 (lane 3) was digested with *Eco*RI, separated by electrophoresis, Southern blotted and hybridised with *gus* sequence. Lane 1: Mixture of pBin19-GTG DNA digested with *Eco*RI and three separate digests of pUCK1 DNA with *Eco*RI, *Hin*dIII or *Eco*RI/*Hin*dIII. Numbers on the left of the autoradiograph are DNA sizes in kb.

internal *Eco*RI fragment.

It was anticipated that the lower size right border fragments identified by Southern hybridisation would be amenable to cloning by inverse PCR.

3.7 Cloning of plant genomic sequence flanking the T-DNA

Several different techniques were used in attempts to clone plant genomic sequences flanking the T-DNA from plants with promoter tagging events.

3.7.1 Inverse PCR

This technique amplifies DNA sequence external to outward facing PCR primers after a ligation step to circularise DNA fragments so that the two PCR primers amplify across the ligation boundary (refer Section 2.7.3).

Primers were designed to allow inverse PCR after either *Hin*dIII or *Pst*I digests with the primer pair RBG/DSO or *Eco*RI digests with primers RBG/DSG (Figure 2).

The methodology (Section 2.7.3) was tested on pGTG. Plasmid DNA was digested with *Sph*I (and *Bgl*II to ensure linearisation of the molecule to allow digestion of *Sph*I) which freed a 4.7 kb *Sph*I fragment across the T-DNA right border. This fragment included the RBG and DSO primer sites, facing outwards, near its two ends. The DNA was treated with ligase in a volume of 1 ml to encourage monomeric circularisation of the DNA fragments, then ethanol precipitated before PCR. The sample was pretreated for 1 minute at 94°C to nick the DNA before 30 PCR cycles of: 15 seconds 94°C, 30 seconds 65°C and 1 second 76°C, in a Hybaid thermocycler with Promega Taq polymerase and the RBG and DSO primers. An aliquot of the reaction was separated by electrophoresis (Figure 24, lane 2) revealing an inverse PCR product of 1 kb as expected.

The protocol was repeated with 5 μg of either E1 or 118 genomic DNA cut with

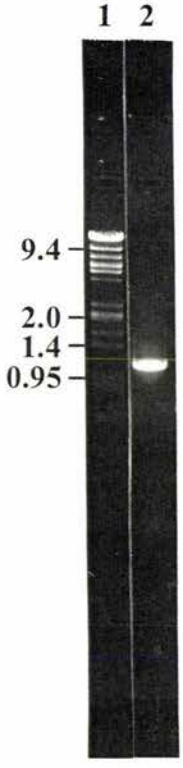


Figure 24. Inverse PCR trial on pGTG.

Plasmid pGTG DNA was *Sph*I digested and ligated (refer text) and amplified with RBG and DSO primers. Lane 1: *Hin*dIII and *Hin*dIII/*Eco*RI digested λ DNA size standards (Section 2.3.12). Lane 2: Inverse PCR treated pGTG DNA.
Numbers on the left of the figure indicate band sizes in kb.

*Hin*dIII. After ligation 1 μg was amplified by PCR. One tenth of the reaction was taken for electrophoresis (Figure 25 **A**) revealing bands of 0.2 kb with the E1 DNA (lane 3) and 0.75 kb with the 118 DNA (lane 4). Control reactions with undigested DNA produced the same non-specific bands with E1, 118 and untransformed plant DNA (lanes 6, 7 and 8). These bands did not match the specific inverse PCR bands.

The protocol was modified to improve the yield of the reaction and to adapt to changes in thermocycler (from Hybaid to Perkin Elmer thermocycler model 480) and Taq polymerase (from Promega to Cetus). A cycle of: 1 minute 94°C, 1 minute 65°C, 5 minutes 72°C was adopted, with a pretreatment of DNA alone (before adding enzyme, buffer and deoxynucleotides) of 5 minutes at 94°C. Promega Taq polymerase gave higher yields with a denaturation temperature of 92°C and in the absence of mineral oil as opposed to 94°C and mineral oil with Cetus Taq polymerase (data not presented, though compare Figure 25 **B** lane 3, including mineral oil, with lane 4, without mineral oil, both using Cetus Taq polymerase; also with Figure 25 **A** lane 3 using Promega Taq polymerase without oil), each with commercially supplied reaction buffer.

Magnesium concentration (1.5 mM in final PCR solution) was not adjusted from the commercial buffers, though the effect of DNA concentration was tested, which would have also titred available magnesium ions in the reaction. A concentration of 1 μg DNA per PCR reaction was found to be optimal under these conditions (data not presented).

The modifications to PCR conditions improved the yield of the major PCR bands (compare Figure 25 **A** lane 3 with **B** lane 4, and **A** lane 4 with **B** lane 6) as well as generating reproducible minor bands.

Increasing the number of cycles from 30 (Figure 25 **C** lanes 5 and 6) to 40 (Figure 25 **C** lanes 3 and 4) also increased the amount of PCR product as seen in reactions (primers RBG/DSG) with bands of 0.75 kb with E1 DNA and 1.2, 0.9 and 0.7 kb with 118 DNA following *Eco*RI treatment. Forty cycles also resulted in another of band of 0.75 kb with 118 DNA (Figure 25 **C** lane 3), though this was not clonable with *Bam*HI and *Eco*RI digests so may not have been a T-DNA specific product (also see

Figure 25. Inverse PCR products following *Hin*dIII, *Eco*RI and *Pst*I digestion of DNA from plants E1 and 118.

A

Lane 1: *Hin*dIII digested λ DNA size standards (Section 2.3.12). Lane 2: *Hae*III digested pUC8 DNA size standards (Section 2.3.12). Remaining lanes are PCR products of: plant E1 DNA digested with *Hin*dIII and ligated (lane 3); plant 118 digested with *Hin*dIII and ligated (lane 4); untransformed tobacco DNA digested with *Hin*dIII and ligated (lane 5); undigested plant E1 DNA (lane 6); undigested plant 118 DNA (lane 7) and undigested untransformed tobacco DNA (lane 8). Primers were RBG/DSO and PCR cycle conditions were 94°C 20 seconds, 65°C 40 seconds, 76°C 1 second for 30 cycles in a Hybaid thermocycler with Promega Taq polymerase and without paraffin oil.

B

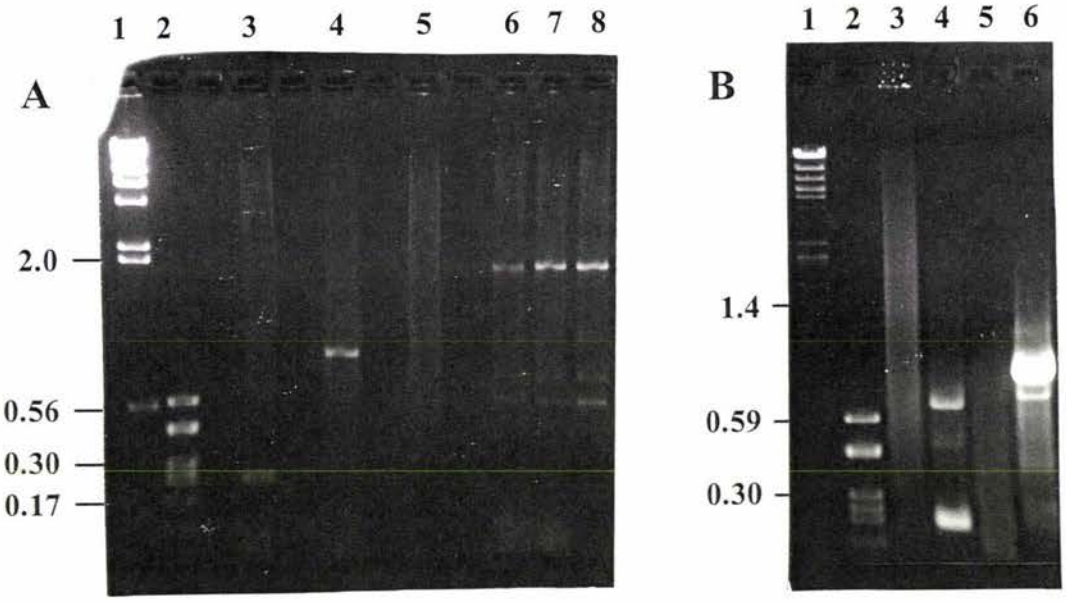
Lane 1: *Hin*dIII digested λ DNA size standards (Section 2.3.12). Lane 2: *Hae*III digested pUC8 DNA size standards (Section 2.3.12). Remaining lanes are PCR products of plant DNA digested with *Hin*dIII and then ligated. PCR primers were RBG/DSO and PCR cycle conditions were 92°C 1 minute, 65°C 1 minute, 72°C 5 minutes for 40 cycles in a Perkin-Elmer thermocycler with Cetus Taq polymerase. Lanes 3 and 4: plant E1 DNA. Lanes 5 and 6: plant 118 DNA. Lanes 3 and 5: PCR with paraffin oil. Lanes 4 and 6: PCR without paraffin oil.

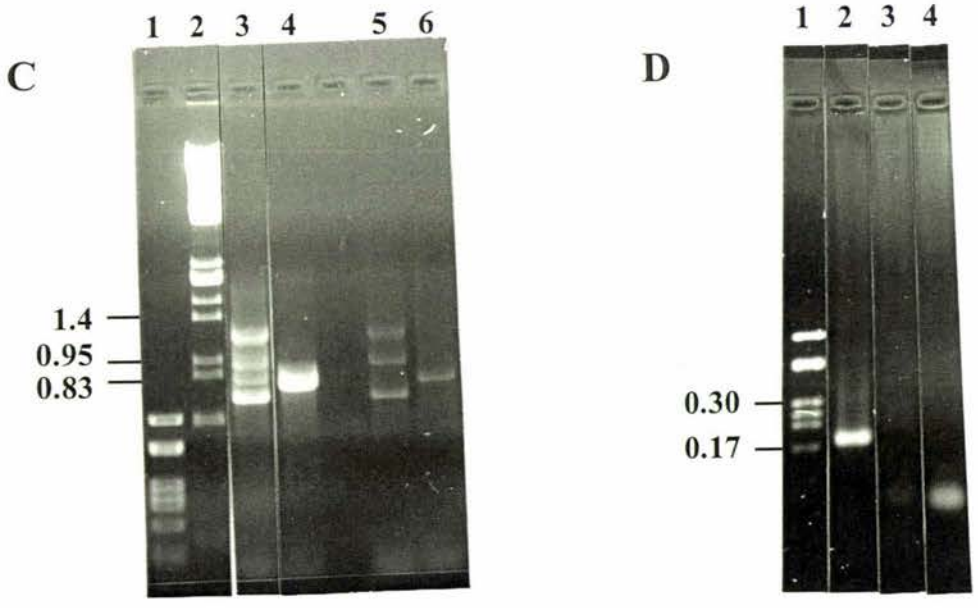
C

Lane 1: *Hae*III digested pUC8 DNA size standards (Section 2.3.12). Lane 2: *Hin*dIII plus *Hin*dIII/*Eco*RI digested λ DNA size standards (Section 2.3.12). Remaining lanes are PCR products of plant DNA digested with *Eco*RI and then ligated. PCR primers were RBG/DSG. Lanes 3 and 5: plant 118 DNA. Lanes 4 and 6: plant E1 DNA. PCR cycle conditions for lanes 5 and 6 were 92°C 1 minute, 65°C 1 minute, 72°C 5 minutes for 40 cycles in a Perkin-Elmer thermocycler with Cetus Taq polymerase and with paraffin oil. PCR conditions were identical for lanes 3 and 4 except that only 30 cycles were conducted.

D

Lane 1: *Hae*III digested pUC8 DNA size standards (Section 2.3.12). Remaining lanes are PCR products with primers RBG/DSO of plant E1 DNA (lane 2), plant 118 DNA (lane 3) and untransformed tobacco DNA (lane 4) following *Pst*I digestion and ligation. PCR cycle conditions were 94°C 1 minute, 65°C 1 minute, 76°C 10 seconds for 40 cycles in a Perkin Elmer thermocycler with Promega Taq polymerase (no oil).


A
1 2 3 4 5 6 7 8
2.0
0.56
0.30
0.17
B
1 2 3 4 5 6
1.4
0.59
0.30


C
1 2 3 4 5 6
1.4
0.95
0.83
D
1 2 3 4
0.30
0.17

Sections 3.7.4 and 3.6) and was only seen sporadically in repeated PCR reactions.

A PCR product (primers RBG/DSO) of 0.18 kb (Figure 25 **D** lane 2) was obtained after *Pst*I treatment of E1 DNA, though no bands were obtained with 118 DNA after *Pst*I treatment (Figure 25 **D** lane 3). The low molecular weight product (Figure 25 **D** lane 4) seen with untransformed tobacco DNA was typical of PCR reactions lacking primer homology within the target DNA. The product is likely due to primer dimerisation.

The right border bands identified by Southern hybridisation (Section 3.6) after *Eco*RI digestion should correspond to the bands generated by inverse PCR reactions, with a difference of 1.9 kb (the *gus* gene present in the Southern fragments but removed by the inverse PCR protocol). For E1, this corresponds to a 0.7 kb fragment that was obtained and cloned by inverse PCR (Figure 25 **B** lane 4). The two larger E1 *Eco*RI fragments would have been too large to amplify under the PCR conditions used. For 118, four *Eco*RI fragments are in the size range amplifiable by PCR, however the sizes, 1.2, 1.1, 0.9 and 0.6 kb, do not exactly match the bands of 1.2, 0.9, 0.75 and 0.7 kb seen after inverse PCR (Figure 25 **C** lane 3). It is possible that small rearrangements within the *gus* sequence could have varied the size of the fragments, or that these bands represent different DNA sequences.

PCR products were subcloned into pUC8 and characterised by sequencing. Thus the PCR reactions products were purified by phenol/chloroform extraction and digested with appropriate enzyme combinations; *Bam*HI cutting at the RBG primer end, *Eco*RI cutting at the end generated by the DSG primer and *Hin*dIII cleaving with primer DSO. For example, PCR products generated with the primer pair DSG/RBG following *Eco*RI treatment (Figure 25 **C** lanes 5 and 6) were digested with *Bam*HI and *Eco*RI (Figure 26). The PCR product bands (Figure 26 lanes 1 and 3, predigestion) had characteristic size decreases following digestion (Figure 26 lanes 2 and 4) as the restriction digests removed the DNA between DSG and the *Eco*RI site (refer Figure 2). This size reduction and the availability of the *Bam*HI site within the RBG primer at the other end of the PCR product were together an indication of the specificity of the reaction to T-DNA derived inverse PCR products. The digested products were ligated into similarly cut pUC8 creating the plasmids pDK1 (118

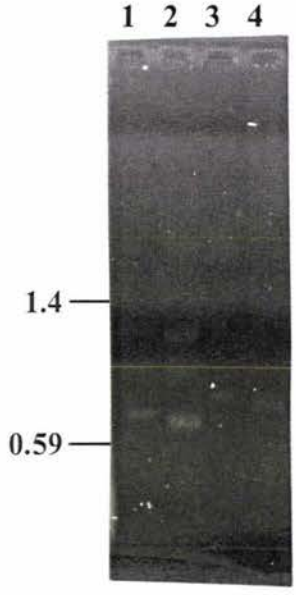


Figure 26. Inverse PCR products from plants E1 and 118 digested with *Eco*RI and *Bam*HI.

Inverse PCR products after *Eco*RI digests of genomic DNA from plants 118 (lane 1 from Figure 25 **C** lane 5) and E1 (lane 2 from Figure 25 **C** lane 6) were digested with both *Eco*RI and *Bam*HI (lanes 3 and 4).

118: lanes 1 and 2.

E1: lanes 3 and 4.

*Hin*dIII inverse PCR, 0.75 kb), pDK2 (118 *Eco*RI inverse PCR, 1.2 kb), pDK3 (118 *Eco*RI 0.9 kb), pDK4 (E1 *Eco*RI 0.7 kb) and pDK5 (E1 *Hin*dIII 0.15 kb).

3.7.2 Analysis of flanking sequence

Inverse PCR products cloned in pUC8 were sequenced from one or both ends with universal and reverse primers. As expected, cloned sequences were about 50 to 100 base pairs shorter than the original inverse PCR products due to loss of a short fragment adjacent to the internally facing T-DNA primer, removed by either *Eco*RI or *Hin*dIII digestion.

Partial sequence from the *Bam*HI end of the 1.2 kb 118 *Eco*RI derived fragment (in pDK2) was found to match the expected T-DNA sequence till the right border and then to match sequence within the 3′ *ocs* region downstream of the *gus* code within the T-DNA (Figure 27 **A**). Hence this clone represented a junction between some form of repeated, and truncated, T-DNA insertion.

Sequence was not obtained using universal primer with the 0.9 kb 118 *Eco*RI clone (in pDK3). It appeared that the primer was recognising two sites within the plasmid. Sequencing instead using RBG as primer generated a short sequence with homology to sequence external to the left border of pBin19 (Figure 27 **B**) which would also be found in a similar location in pBin19-GTG. Sequencing with reverse primer from the opposite end of the clone found sequence matching the α-complementation fragment of β-galactosidase (Figure 27 **B**). This is present near the left border of the T-DNA. Hence this clone also represents the junction of some form of T-DNA repeat, in this case a T-DNA which extends beyond the normal left border.

The *Hin*dIII and *Eco*RI inverse PCR clones (in pDK5 and pDK4 respectively) from E1 were found to overlap. The *Hin*dIII derived sequence extended 0.15 kb from the *Bam*HI site within the RBG primer, while the *Eco*RI derived sequence continued, creating a fragment of 0.7 kb. This sequence also contained an internal *Pst*I site 0.1 kb from the *Bam*HI site, so the E1 *Pst*I inverse PCR product of 0.18 kb was not individually subcloned. Partial sequence of this 0.7 kb clone (Figure 27 **D**) and full

A

```
  1  CGTCAATGGG CTTCCaTtGT CAATAatGAA ACCACGGCTC aAGACTTACC
 51  AGCGGCAGGT AATTTGTAGT ACATCCAACA CTGatagttt
```

B

```
     gaattc ... (few bases) ...
  1  cCGTTTTACA ACGTCGTGAC TGGGAAAACC CTGGCGTcAC CCAACTTAAT
 51  CGCCTTGCAG CACATCCCCC TTcgGCCAGC TGGCGTAATA GCGAAGAGGC
101  CTGCACCGAT CGCCCTTCCC AgCAGTTGCG CA
                  ....(many bases)....
751  TAAGGCGGCA GACTTTGCTC ATGTTACCGA TGCTATTCGG AAGAgCGGCA
801  ACTAAGCTGC CGGGTTTGAA ACACGGATGA TCTCGCGGAG GGTAGCAatg
851  ttgtaacgat gacaggaag ... (few bases)...ggatcc
```

Figure 27. DNA sequence from inverse PCR cloning isolates.

Matches to the inverse PCR primer RBG and the sequence upstream of the primer to the T-DNA right border are underlined.

A Sequence from near *Eco*RI end of the 1.2 kb *Eco*RI derived 118 isolate. Sequence matches to the octopine synthase 3′ region are capitalised.

B Sequence from the 0.9 kb *Eco*RI derived 118 isolate. Sequence matches to pBin19 capitalised. Sequence capitalised on the upper half of the diagram matches the *lacZ* α-complementation region within the T-DNA and sequence capitalised on the lower half of the diagram matches pBin19 sequence external to the left border.

C Sequence from the 0.75 kb *Hin*dIII derived 118 isolate.

D Sequence from the 0.7 kb *Eco*RI derived E1 isolate. The *Hin*dIII and *Pst*I restriction sites are italicised and labelled.

C

```
   1  aagctttatt aaatGTGatt ttgctaaata atttatcaat gtaattacta
  51  ctttttgcag tgtttacgcg ttgaataagg tgtagaaaaa agtttataaa
 101  aagttgaaat aagataaaaa agaatagccg tgagagagtc aaaacaacga
 151  gacctaattt ctggctgcac cccattaaca tcttttcaaa gaaactacag
 201  taactttgat ttttgaccca tttccttttc tgaccataca aaaatatgaa
 251  gattattcca tgtccaattt catggctaat ttccaaaact accaatttct
 301  aggtttttcg ttcaagaagt cccataattt ctacaaattt tcatgtttaa
 351  atccttatac cactatgtat ttaactcaca atgggggaat aacttacctt
 401  gccattgatg atgaaaccgg tcctcttgaa gctctccaag atcgctctca
 451  aaccaaaaga gaatgaaagt aaatgggcca atcccgttc ttaaaagaac
 501  acactgcccc gcgacctctt gcacccataa ggtacgcctg gtcgagcagg
 551  cgacgcggat cgacgtccag ngcacgagag aatggtccgc catctGTGta
 601  atgcgtctag gatcgagaca aaggtcgtct ctgtcagccg gatcgtagaa
 651  catattgcgg cgccaaagcc cacttcaccg accagaccga cgggatctat
 701  caccagccag ccgcgaaaca ctgatagttt aaaccgaagg cgggaaacga
 751  caatctgatc ccggggatcc
```

D

```
   1  attgcaGTGg catgcataac atgataacat gcacaagagt tgaaggcaga
  51  gaggaaacta cagaaggaga gattcagcta aatggagaag caacagagag
 101  tgtatgatgt gatattgttt tttggtatca tagacatact tcaagactat
 151  gacattacaa agaagcttga gcatgcatac aagtctatgc aatGTGatcc
                 HindIII
 201  cgactccatc tctgcagtcg atccaaaggc gtactcaagg agttttcGTG
                  PstI
 251  attacatttt caaacactga tagtttaaac cgaaggcggg aaacgacaat
 301  ctgatcccgg ggatcc
```

sequence of the 0.75 kb 118 *Hin*dIII clone (in pDK1, Figure 27 **C**) revealed no homology to T-DNA sequences beyond a short match from the *Bam*HI site to the right border. They were also found to have no significant open reading frames nor homology to any other sequence within the Genbank database.

In order to clone the additional major bands seen after inverse PCR reactions (0.7 and 0.75 kb after *Eco*RI treatment of 118 DNA) and other minor bands, the PCR reaction products were re-amplified in a second round of PCR using the same primer combinations. Subsequent cloning only revealed sequences which were limited in homology to the RBG primer sequence itself. These products were presumably the result of mispriming during PCR. As well, in order to ensure cloning of inverse PCR products that may have had an internal *Bam*HI site, the double digested inverse PCR products were ligated into *Bam*HI linearised pUC8. However, only doublets were obtained of previously cloned band sizes.

3.7.3 Plasmid rescue

An alternative approach to clone plant genomic sequence flanking T-DNA insertions is plasmid rescue.

Initially 5 µg of tobacco genomic DNA was chosen as a suitable quantity for plasmid rescue. Given a haploid genome of 6.4 x 10^9 base pairs (hence mass of 6.9 picograms), 5 µg of DNA will contain 360 000 copies of a T-DNA represented once in the diploid genome.

DNA from plant E1, either *Eco*RI (for rescue of a potential T-DNA and left border flanking region) or *Bgl*II (for rescue of a potential T-DNA and right border flanking region) digested, was self-ligated in a volume of 100 µl and 1 µg was used to transform DH5α competent ($CaCl_2$) cells which were plated onto LB 100 mg l^{-1} ampicillin. A large number of transformants resulted, 95 from the *Eco*RI treatment and 54 from the *Bgl*II treatment. These were screened and found to consist of two different size classes, each with a consistent restriction pattern that did not match the expected T-DNA restriction pattern. Both classes of transformant were found in each

initial restriction enzyme treatment.

The plasmid rescue technique was refined to prevent false positive transformants resulting from contaminating plasmid DNA and to increase the transformation efficiency, as detailed in Section 2.7.4.

Transformants were screened on XGal/IPTG plates. The presence of any blue colonies asserted that further precautions were necessary. However separate stocks of restriction endonucleases and ligase were not necessarily available, and the same micropipettes were used as for other work. Control treatments with, for example, untransformed tobacco genomic DNA either untreated, ligated only or *Eco*RI digested and ligated, were conducted to eliminate sources of contaminating DNA with some success. However transformants arising from sources other than the T-DNA were never completely eliminated.

E.coli strain DH5αMCR was used as the host strain in order to avoid restriction of methylated plant genomic DNA (Grant *et al.*, 1990). When electroporation was adopted, DH10b was used as the host strain. This strain has a high transformation efficiency with electroporation and also avoids the sensitivity of electroporated DNA to restriction methylation by the host (Grant *et al.*, 1990; Lorow and Jessee, 1990). Electroporation using a Biorad Gene Pulser and Pulse Controller was found to markedly improve transformation efficiencies from a typical $2x10^5$ (using $CaCl_2$ competent cells) to $1x10^9$ colonies per μg of pUC8 DNA. The presence of 5 μg of plant genomic DNA was found to inhibit the electroporation efficiency of 20 pg pUC8 DNA by 60%. This level of inhibition was acceptable and this quantity of plant genomic DNA was adopted as standard for electroporation. Later, electroporation was carried out on a BRL Cell-Porator with Voltage Booster. A single trial, with duplicate samples for each machine, compared the transformation efficiencies of the two electroporation machines. The BRL Cell-Porator, with a field strength of 16 kV cm^{-1}, was found to have 120% more colonies than the Biorad Gene-Pulser, with a field strength of 12.5 kV cm^{-1}.

Finally 20 μg samples of genomic DNA from plants E1 and 118 were digested with either *Hin*dIII or *Eco*RI and ligated with 12 units of ligase in a volume of 400 μl.

After electroporation no transformants were seen from the 'E1 *Hin*dIII' treatment, but 18 from 'E1 *Eco*RI', 2 from '118 *Hin*dIII' and 30 from '118 *Eco*RI' treatments. The '118 *Eco*RI' treatment should facilitate rescue of an internal T-DNA fragment, which is present in the pBin19-GTG T-DNA, but not in the pGTG T-DNA, though this would be impossible to distinguish from contamination of the plant DNA with pBin19-GTG plasmid as their sequence would be identical.

All colonies were screened on XGal/IPTG and a selection of their plasmids were restriction mapped.

Three colonies were found to be blue. One of these was from the '118 *Eco*RI' treatment, as well as both the '118 *Hin*dIII' transformants, though one of these was a very faint blue colour. This isolate yielded a 2.7 kb plasmid with no *Pst*I site and single *Eco*RI and *Hin*dIII sites 230 base pairs apart. The other blue isolate from the same treatment contained a 2.2 kb plasmid with no *Eco*RI or *Hin*dIII sites but one *Pst*I site.

Two colonies from the '118 *Eco*RI' treatment contained large plasmids of 7.5 kb, with no *Hin*dIII, one *Eco*RI and three *Pst*I sites. These did not match the expected internal T-DNA fragment of 3 kb with one *Eco*RI and one *Pst*I site. Hence these clones could be more contaminants.

Two isolates from the 'E1 *Eco*RI' treatment contained plasmids of 4.5 kb. One, named pDK6, was restriction mapped and compared to the left border region of pGTG (Figure 28). The maps match from the *Eco*RI site, through *Pst*I and *Pvu*II sites to the *Bgl*II sites 1 kb inside pGTG's T-DNA left border. A *Cla*I site 0.5 kb from the left border is missing from the plasmid rescue clone. It is possible that this clone is an artifact or contaminant, but also possible that it is an actual plasmid rescue of a partially truncated T-DNA and the plant genomic sequence flanking its left edge. To confirm this it would be necessary to subclone the potential plant genomic portion of the clone and use it as a probe against tobacco genomic DNA. This would give a different hybridisation band between untransformed tobacco and transgenic E1 due to the T-DNA insertion, if the clone is a result of plasmid rescue. DNA sequence alone would only be sufficient to disprove the possibility of successful plasmid rescue by

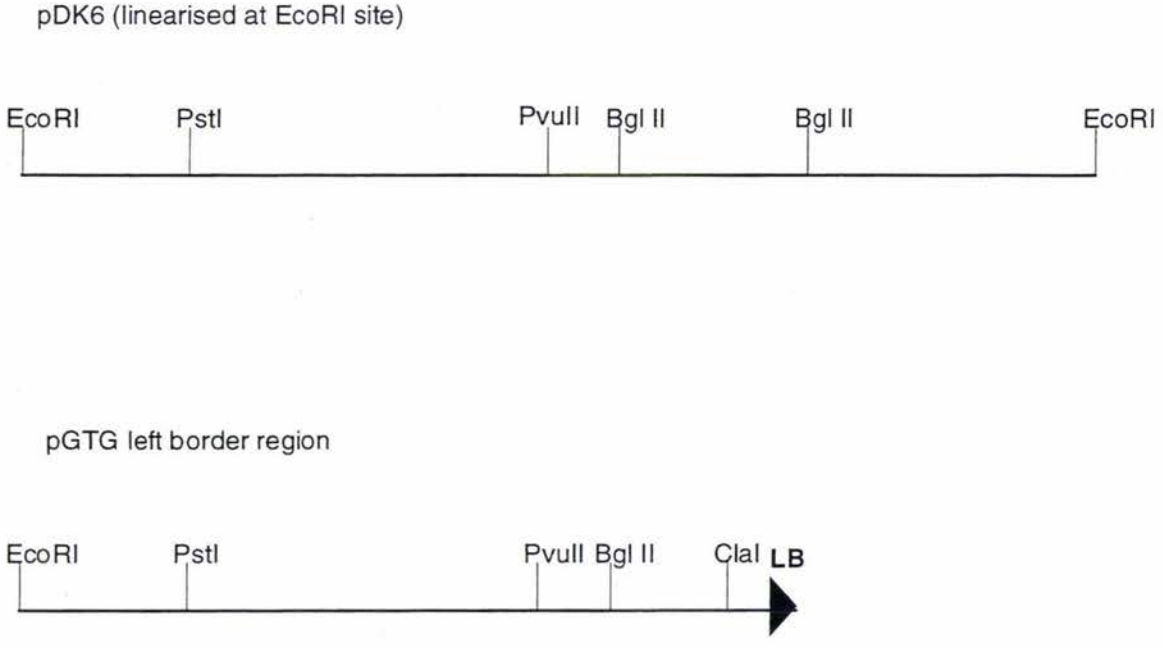


Figure 28. Comparison of restriction maps of pGTG T-DNA left border region and plasmid rescue clone pDK6.

Abbreviation: LB, left border of T-DNA.

showing lack of homology to T-DNA sequence, or homology to another source of DNA.

Due to the complexities of this approach experiments were halted at this point. Further work was required, especially mapping and DNA sequencing of '118 *Eco*RI' clones to ascertain if the bulk of these clones were, as anticipated, possible plasmid rescues of an internal T-DNA fragment. Also, sequencing other clones could have identified the source of the contaminants.

3.7.4 Single sided ligation mediated PCR

Single sided ligation mediated PCR was chosen as an alternative method to clone plant genomic sequence flanking the T-DNA right borders. Like inverse PCR, it should allow direct amplification of these sequences for cloning, but a greater range of restriction enzymes are available for generating the plant genomic end of the fragment. Hence fragments of an amplifiable size could be expected from more of the T-DNA insertions by using additional restriction enzymes.

First, this procedure (Section 2.7.5) was trialled (Figure 29) on various digests of both pGTG and pBin19-GTG designed to yield PCR products of lengths from 190 base pairs to 4.9 kb.

The enzyme *Hin*dIII was included in the *Sph*I digest to linearise the plasmid as *Sph*I had previously been noted to be inefficient at cutting circular plasmid DNA. The smallest expected product, a 190 base pair fragment from the RBG primer to the *Sph*I site immediately upstream of the right border was generated, as well as an unexpected 1.2 kb fragment in both pGTG and pBin19-GTG reactions (Figure 29, lanes 4 and 7). This could be interpreted as resulting from ligation of the 1.0 kb *Sph*I-*Hin*dIII fragment internal to the T-DNA in both plasmids to the 0.2 kb right border fragment, and amplification of this chimaeric molecule. A *Xho*I digest of pGTG yielded the expected 1.2 kb fragment (lane 6), but *Kpn*I and *Bgl*II digests did not generate 3.6 kb and 4.9 kb fragments (lanes 5 and 3 respectively).

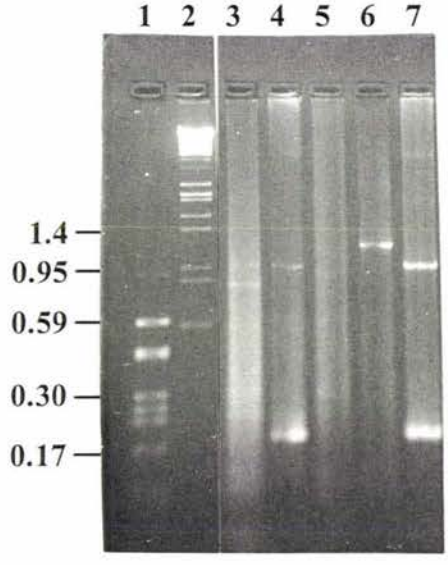


Figure 29. Single sided ligation mediated PCR trial on plasmid DNA.

Single sided ligation mediated PCR (Section 2.7.5) was trialled on plasmid DNA following *Bgl*II digestion of pGTG DNA (lane 3), *Sph*I/*Hin*dIII digestion of pGTG DNA (lane 4), *Kpn*I digestion of pGTG DNA (lane 5), *Xho*I digestion of pGTG DNA (lane 6) and *Sph*I/*Hin*dIII digestion of pBin19-GTG DNA (lane 7).

Lane 1: *Hae*III digested pUC8 DNA size markers (Section 2.3.12).

Lane 2: *Hin*dIII plus *Hin*dIII/*Eco*RI digested λ DNA size markers (Section 2.3.12).

Next, 3 μg of genomic DNA from plants E1 and 118 digested with either *Eco*RI or *Hin*dIII was analysed using the same procedure, except that 30 PCR cycles were run instead of 40. Samples of the PCR reactions were run on a gel (Figure 30 **A**). Non-specific banding overlaying a smear of DNA was apparent in all three *Eco*RI lanes (Figure 30 **A** lanes 5, 7 and 9) including the untransformed tobacco genomic DNA (lane 9). A single band appeared at 0.2 kb in the E1 *Hin*dIII treatment (Figure 30 **A** lane 4), below a smear also appearing in all three *Hin*dIII samples (Figure 30 **A** lanes 4, 6 and 8) including the untransformed tobacco control (lane 8).

The gel was blotted and probed with the 0.75 kb *Bam*HI-*Hin*dIII fragment of pDK1 (plant 118 inverse PCR isolate) eluted onto DEAE paper. This fragment will have homology to all putative single sided ligation mediated PCR products from the *Bam*HI sited at the 5′ end of the RBG primer to the right border itself, a sequence of 53 base pairs.

Specific bands were seen (Figure 30 **B** lanes 4 to 7) in all four transgenic DNA samples whilst the untransformed lanes were clear (Figure 30 **B** lanes 8 and 9). The marker lane with HaeIII digested pUC8 also hybridised to the probe (Figure 30 **B** lane 1) indicating that vector DNA had also been purified with the 118 subclone.

A band of 0.4 kb was found in all four transgenic DNA samples (Figure 30 **B** lanes 4 to 7). This band and the 0.3 kb band below it found in 118 lanes (Figure 30 **B** lanes 6 and 7) have no corresponding bands in inverse PCR reactions. On the other hand, the 0.6 and 0.2 kb bands (lane 4) could correspond to the two most intense E1 *Hin*dIII inverse PCR bands (Figure 25 **B** lane 4). Similarly, the 0.7 and 0.6 kb bands (Figure 30 **B** lane 6) could correspond to slightly larger (due to different primers and methodologies between inverse PCR and single sided ligation mediated PCR) bands of 0.75 and 0.65 kb *Hin*dIII derived inverse PCR products (Figure 25 **B** lane 6) and likewise the 0.85 and 0.65 kb bands (Figure 30 **B** lane 7) might correspond to 0.9 and perhaps 0.8 kb *Eco*RI derived inverse PCR products (Figure 25 **C** lane 3).

This, at least partial, correspondence between the two PCR cloning methods indicated that the bands were less likely to be PCR artifacts as different primer combinations were utilised.

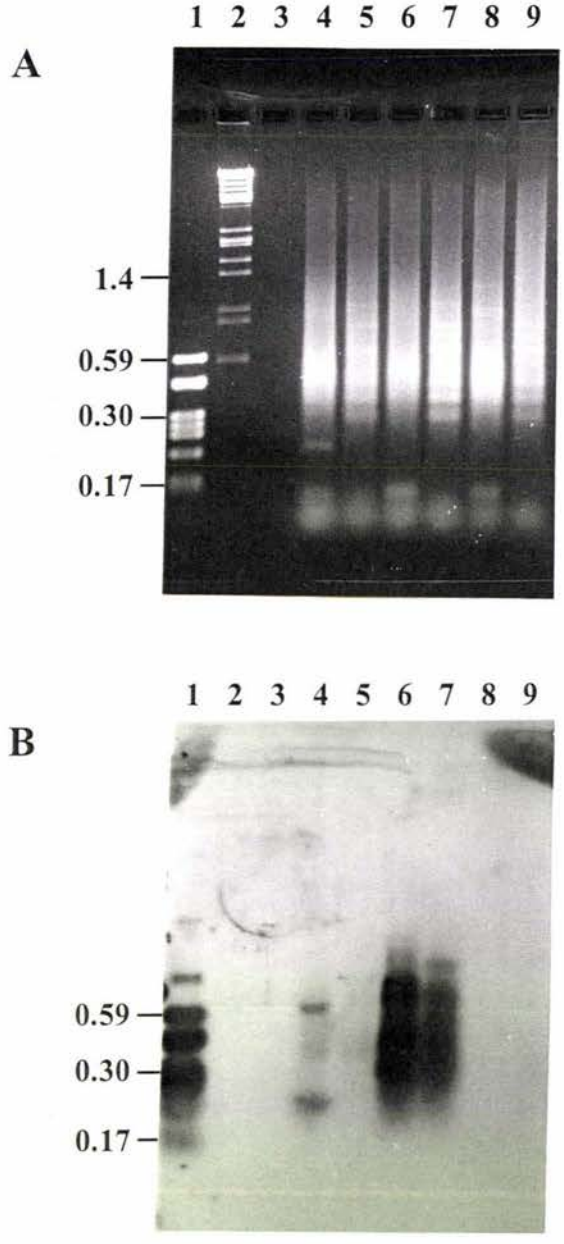


Figure 30. Single sided ligation mediated PCR from plant genomic DNA.

A Single sided ligation mediated PCR (Section 2.7.5) was conducted on genomic DNA from plant E1 (lanes 4 and 5), plant 118 (lanes 6 and 7) and an untransformed plant (lanes 8 and 9). DNA was first digested with *Hin*dIII (lanes 4, 6 and 7) or *Eco*RI (lanes 5, 7 and 8). Lane 1: *Hae*III digested pUC8 DNA size markers (Section 2.3.12). Lane 2: *Hin*dIII plus *Hin*dIII/*Eco*RI digested λ DNA size markers (Section 2.3.12). Lane 3: Blank. Numbers on the left of the photograph are DNA sizes in kb.

B Autoradiograph of **A** after Southern blotting and hybridisation with 0.75 kb *Bam*HI/*Hin*dIII fragment from pDK1.

3.8 Rationale for further experiments

Transformation of the promoter tagging vectors pGTG and pBin19-GTG into tobacco had resulted in a high frequency (37% of transgenic plants) of *gus* activation. However, analysis of two plants (E1 and 118) selected for their root specific GUS activity revealed multiple T-DNA copies, at least of the right border/*gus* fragment, as seen by Southern hybridisation, inverse PCR and single sided ligation mediated PCR. Segregation on kanamycin selection suggested multiple copies of the *npt*II gene were also present at least in plant 118. In progeny, *gus* expression was only seen in a brief window after 5-azacytidine treatment, and then only in plant 118 seedlings. Because of the number of T-DNA copies and the potential difficulty in segregating them by outcrossing, further efforts were not made to clone the tagged promoters in these two transgenics.

Two significant questions were raised in the previous results. One, whether the average T-DNA insertion frequency with these promoter tagging vectors was as high overall as that observed in the two plants analysed. This could be assessed by analysing a larger sample of transgenic plants for T-DNA copy number. Two, whether the plants observed with a low level of GUS activity within root material were truly promoter tagged or merely false positives from the fluorometric screening method due to high background fluorescence of tobacco root material. Histochemical screening for GUS activity would provide an alternative method of identifying *gus* activation, as well as generating finer details on patterns of gene expression.

4.0 Results II

4.1 Histochemical screening of tissue culture roots and shoots

A second population of transgenic plants were generated, transformed with pBin19-GTG (Section 2.4.5). Transgenic tobacco plants in tissue culture growing in individual pottles were analysed at approximately similar stages of growth, about 2 weeks after cut shoots rooted on NicIII media containing kanamycin and cefotaxime. An entire young leaf including petiole, a piece of older leaf and several lengths of root including root tips and sites of lateral root initiation were histochemically assayed for GUS activity (Section 2.5.1).

Two sets of plants were screened. The "A" series of plants were transformants of *Nicotiana tabacum* DSIR line KKD. Only plants that were found to be GUS positive from tissue culture screening were kept for further analysis. The "W" series of plants were transformants of *Nicotiana tabacum* variety Wisconsin 38. All of these plants were kept for further analysis. In both series of plants regenerants from the same leaf piece were given the same first number with a second number to identify different isolates (for example A series plants A1.1, A1.3 and A1.4 were all isolated from one leaf piece, different to that of A18.1, or to W series plant W1.1) so that any shoots arising from the same transformation event could be identified by Southern hybridisation (Section 4.4).

The results of the screening are summarised in Table 9 with more detailed information below (Section 4.2).

4.2 Individual patterns of gene expression revealed by promoter tagging.

The most striking observation to be made from the collection of *gus* expression patterns in the A and W series transformants (Table 10) is the diversity found.

A number of plants had variable staining after repeated testing. For example, A9.2 was tested twice and found on both occasions to stain with XGluc in various shoot

Table 9. Histochemical screening for GUS activity of transgenic *Nicotiana tabacum*.

Transformant series	Number of plants screened	GUS positive			Percentage GUS positive
		shoot only	shoot and root	root only	
A[1]	65	9	4	2	23%
W[2]	82	10	4	4	22%

[1] A series plants are DSIR line KKD transformed with pBin19-GTG.

[2] W series plants are Wisconsin 38 transformed with pBin19-GTG.

Table 10. Distribution of GUS staining in plant organs.

plant	root			shoot	leaf				floral organs
	root tip	lateral initials	other root	shoot apex	young lamina	young petiole	mature lamina	mature petiole	
A1.1	+++[1]	+++	++[2,3]	+	+	+			base of ovary/receptacle +[2], stigma +[2]
A1.4					+	++			sepal +
A9.2		+[2]			++	+++	+	+	nd
A10.2	++								nd
A14.3				+	+		+		
A18.1				++					
A18.2				+					nd
A20.1	+[2]							+[2]	base of ovary/receptacle ++
A26.6	+++	+++	+++	+++	++	++	++	++	nectaries +++, ovary ++, receptacle ++
A29.2				+	+				ovary +
A35.1				++	+	++	++	++	nd
A41.1		++	+[3]	+++		+			
A53.1				++	+	+	+	+	

[1] +++: strong staining, ++: moderate staining, +: weak staining, blank: no staining, nd: not determined. [2] variable [3] vascular tissue.

Table 10 continued.

plant	root			shoot	leaf				floral organs
	root tip	lateral initials	other root	shoot apex	young lamina	young petiole	mature lamina	mature petiole	
A82.1				+					nd
A88.1	++		+[4]	++	+	+	+[6]	+	stigma +, style +
W1.1				+	+				nd
W2.5							+		
W2.6				+	+				
W5.3							+		
W5.5					+	++			
W6.4							+	+	nd
W8.4							+	+	
W9.1				+	+	+		+	nd
W11.1									immature anthers +
W23.2	++	++	++	+++	+++	+++	+++	+++	nd
W24.1	+[5]								

[4] vascular tissue in older roots. [5] and also expression several millimetres behind root tip. [6] midrib

Table 10 continued.

plant	root			shoot	leaf				floral organs
	root tip	lateral initials	other root	shoot apex	young lamina	young petiole	mature lamina	mature petiole	
W26.1	++			+					
W40.1							+[6]	+	receptacle ++
W41.1					+[6]	++	+	++	receptacle/nectaries/base of ovary ++, stigma ++
W42.1	++[2]								ovary +, receptacle +++
W43.2	+	+	+[2,3]				+	+	ovary +, nectaries +, petal +[8]
W43.6			++[7]						base of ovary/receptacle +[2], anther +[2]
W46.3	+++[5]								
W47.2	++	++	++	+++	+++	+++	+++	+++	nd
W48.5									immature anthers +
W50.2					+	+			nd
W53.1									style +

[2] variable. [3] vascular tissue. [5] and also several millimetres behind root tip. [6] midrib only. [7] expression only in root with dead root tip.
[8] zone in corolla tube where colour changes from white to pink

tissues and also in lateral root initials. Months later the plant was retested twice and while the shoot staining remained unchanged, the root staining was absent (Figure 31 **A**). In most cases, the variable expression was seen as a loss of previously seen staining, though A20.1 lost staining in mature leaves and gained it in root tips.

While A1.1 lost staining in floral material (Table 10), it also had variable intensity staining in root tips (Figure 31 **B** and **C**) and in longer sections of root material (Figure 31 **D**). Other plants such as A26.6 were uniformly strong in staining intensity (Figure 31 **M**). It is interesting to compare the root positive plants which frequently expressed in root tips but gave dissimilar patterns of expression at lateral primordia. Plant A1.1 stained strongly in lateral initials (Figure 31 **E**), though this expression grew out with the young root (Figure 31 **F**) compared to expression seen remaining at the branch points with plant 118 observed earlier (Figure 18 **A**). Other root tip expressers did not stain in lateral root primordia. The young lateral root tips of A10.2 and A88.1 only stained after breaking the surface of the parent root, while lateral roots of W46.3 stained just before breaking the root surface.

While A1.1 had strong root expression, shoot material also stained. Staining was not found in the shoot apical meristem, but there was some staining in leaf primordia just below the apex (Figure 31 **G**), as well as weakly in young leaves. One shoot pattern common to some plants was strong GUS activity in the shoot apex with less activity in leaf material. For example A18.1 stained only in the shoot apex (Figure 31 **H**) while A35.1 stained in the apex and in other leaf material (Figure 31 **I**). But other shoot positive plants did not stain in the shoot apex, for example A1.4 with staining only in immature petioles and less in immature leaf lamina (Figure 31 **J**).

Three plants, A26.6, W23.2 and W47.2, had staining throughout tissue culture shoot and root material. Only A26.6 flowers were tested and this revealed that A26.6 *gus* expression was not constitutive, being present only in ovary, nectary and receptacle tissue (Section 4.3) in the flowers at the time of anthesis and not in other floral organs or in flower buds or seed pods.

Distinct patterns of extremely limited staining were indicative that further tagged promoters might be found if more intensive screening were conducted. For instance,

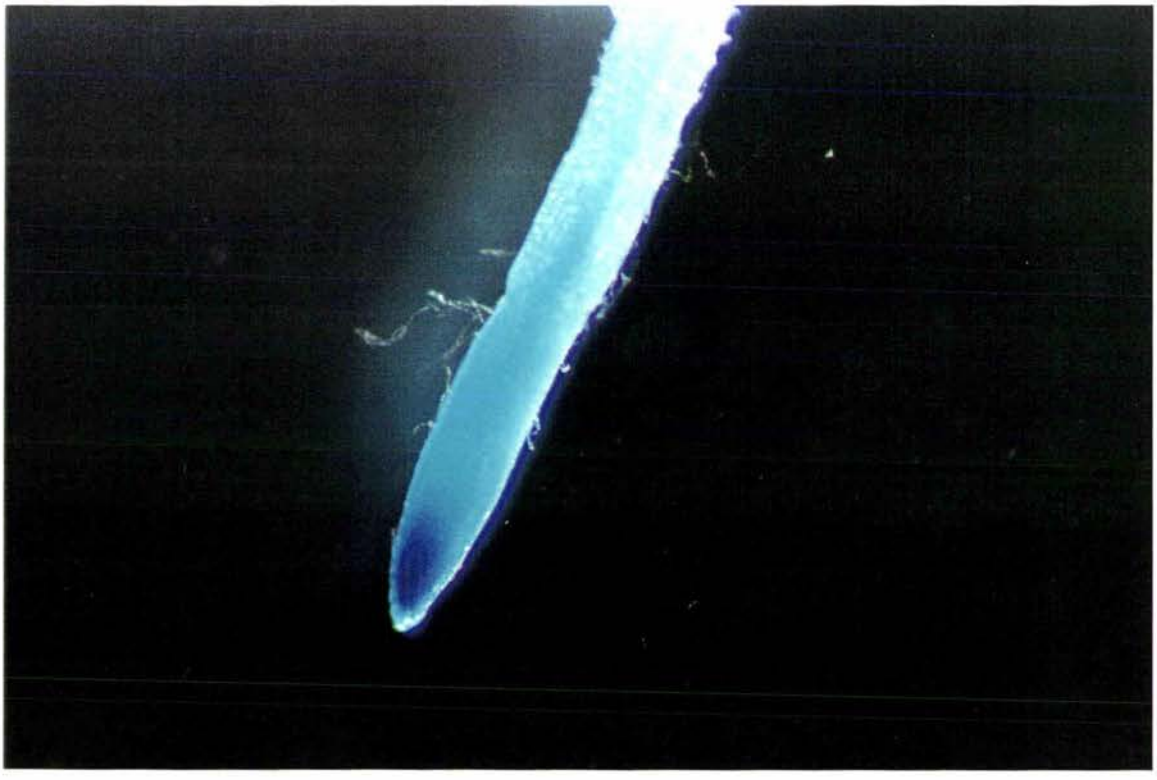
Figure 31. Histochemical GUS detection in diverse organs of transgenic tobacco.

Shoot and root material were sampled from plants growing in tissue culture and stained with XGluc (Section 2.5.1).

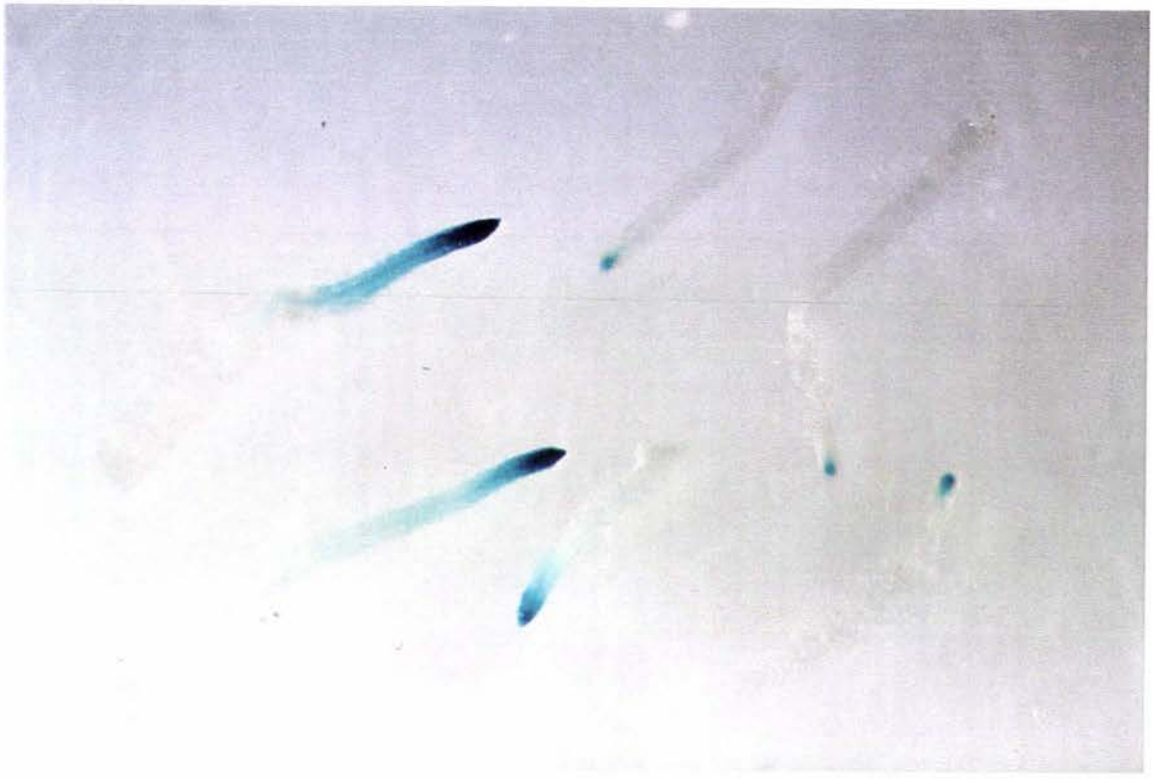
A	A9.2. Clockwise from top: young leaf, roots, segment of mature leaf, petiole from mature leaf
B	A1.1. Root tip.
C	A1.1. Root tips.
D	A1.1. Roots.
E	A1.1. Root with lateral root primordia.
F	A1.1. Lateral root and branch point.
G	A1.1. Shoot apex.
H	A18.1. Shoot apex.
I	A35.1. Shoot apex and leaf material.
J	A1.4. Immature leaf.
K	A26.6. Mature leaf.
L	A26.6. Shoot apex and leaf material.
M	A26.6. Root with lateral root primordia.

A

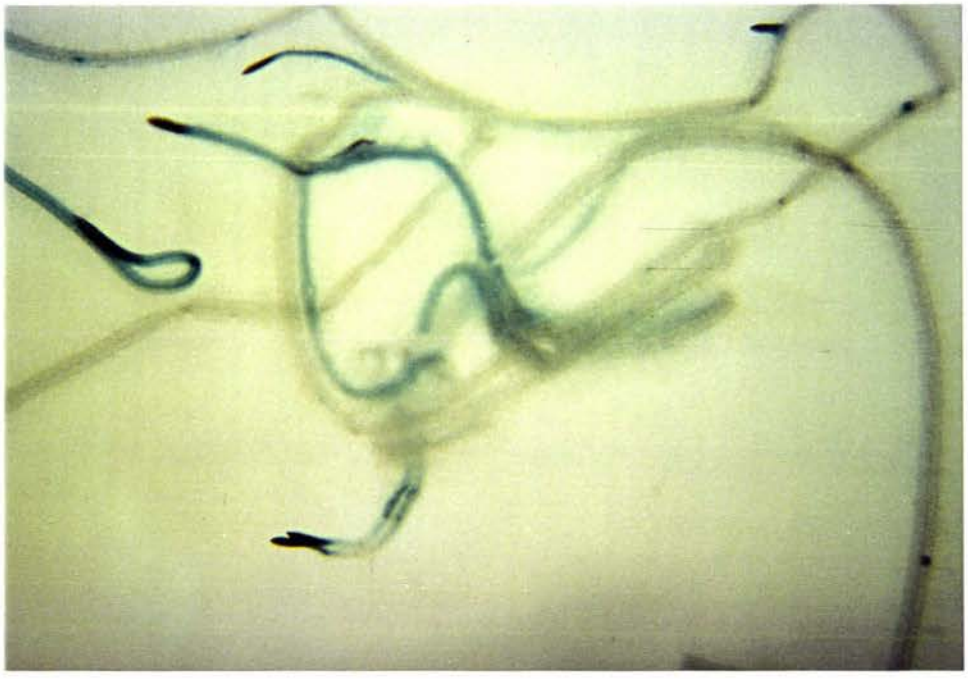
B



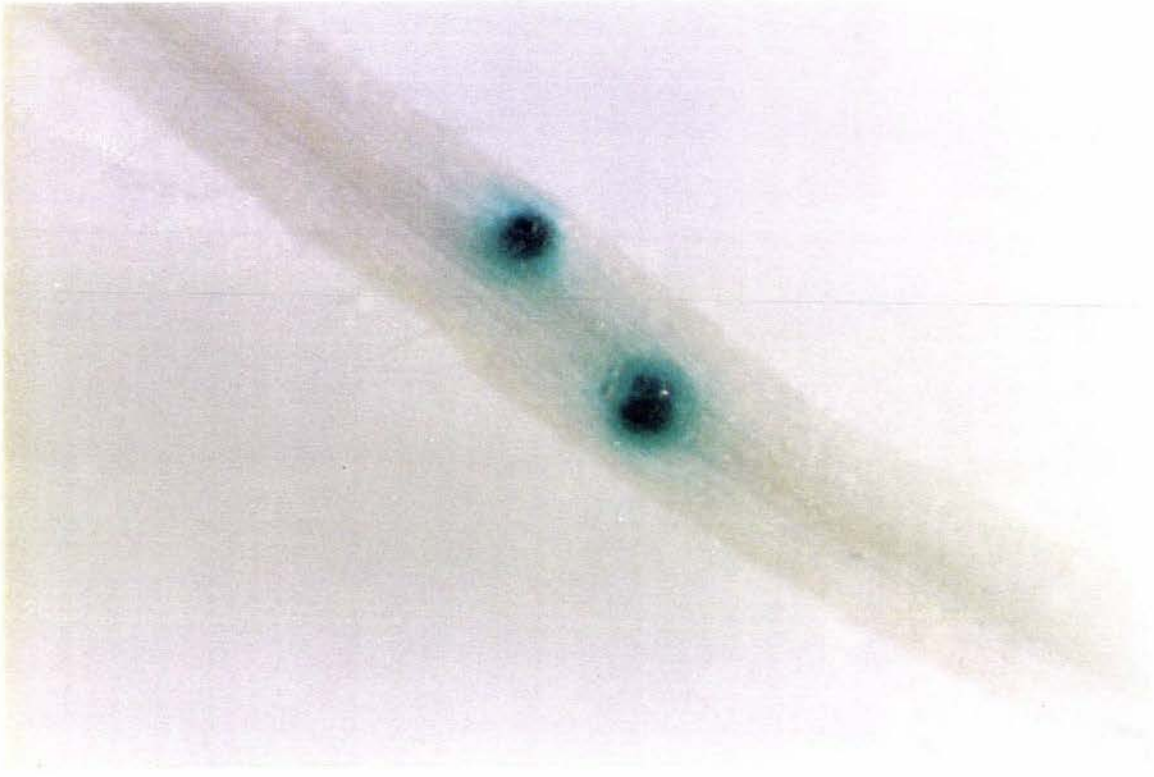
C



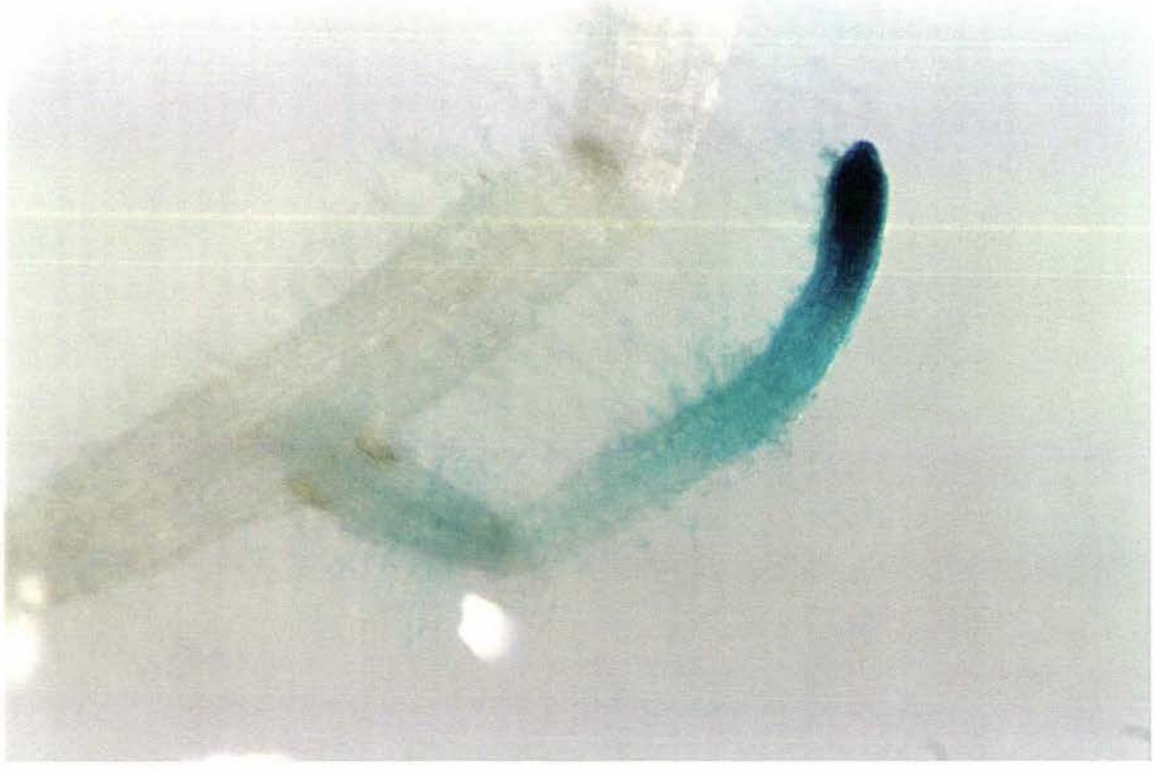
D



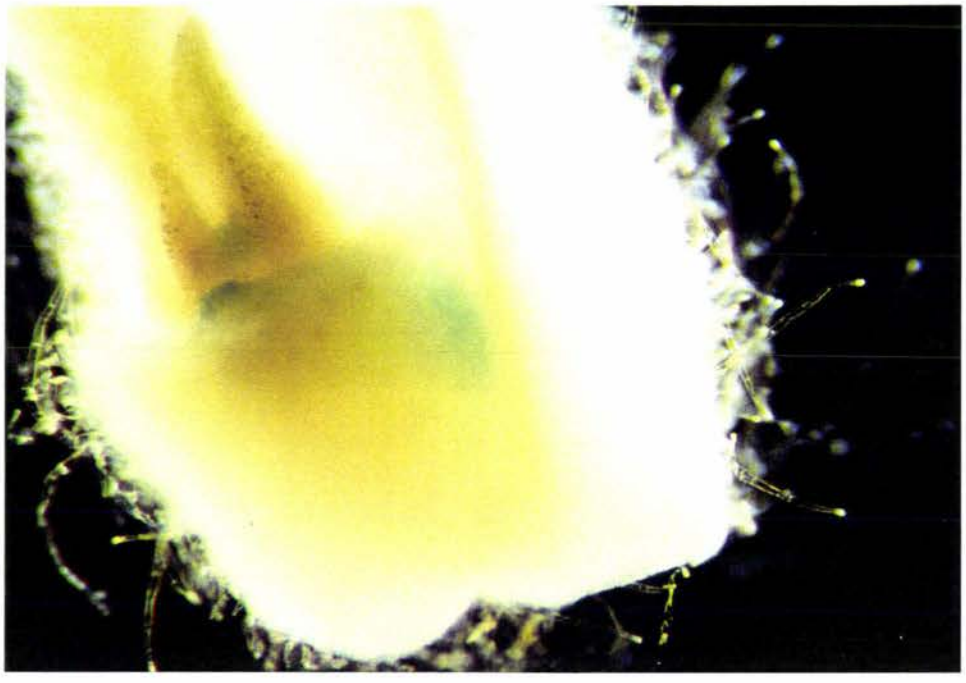
E



F



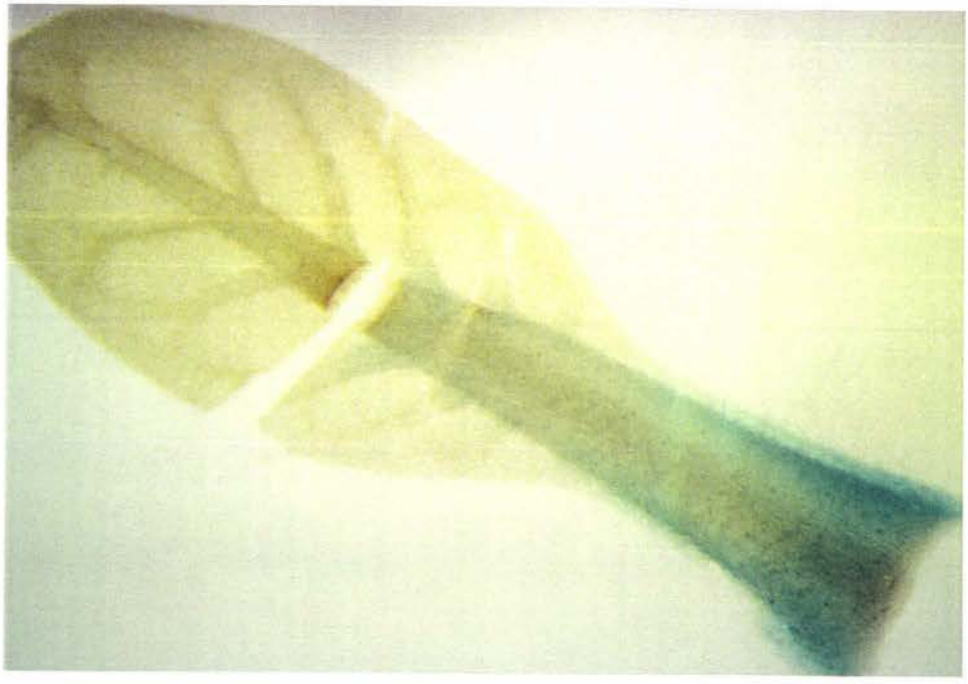
G



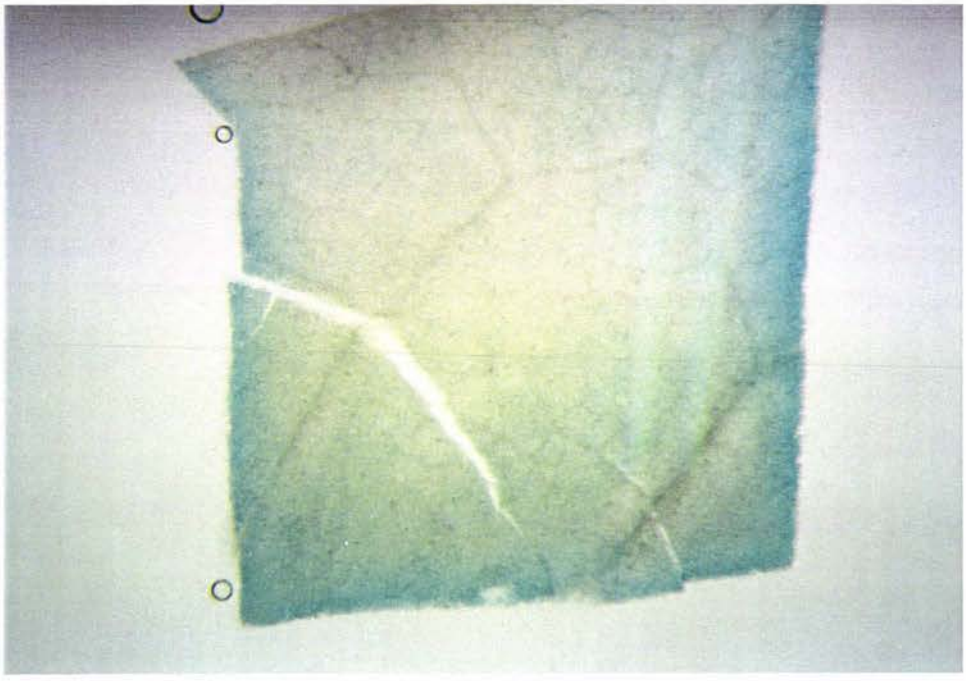
H

I

J

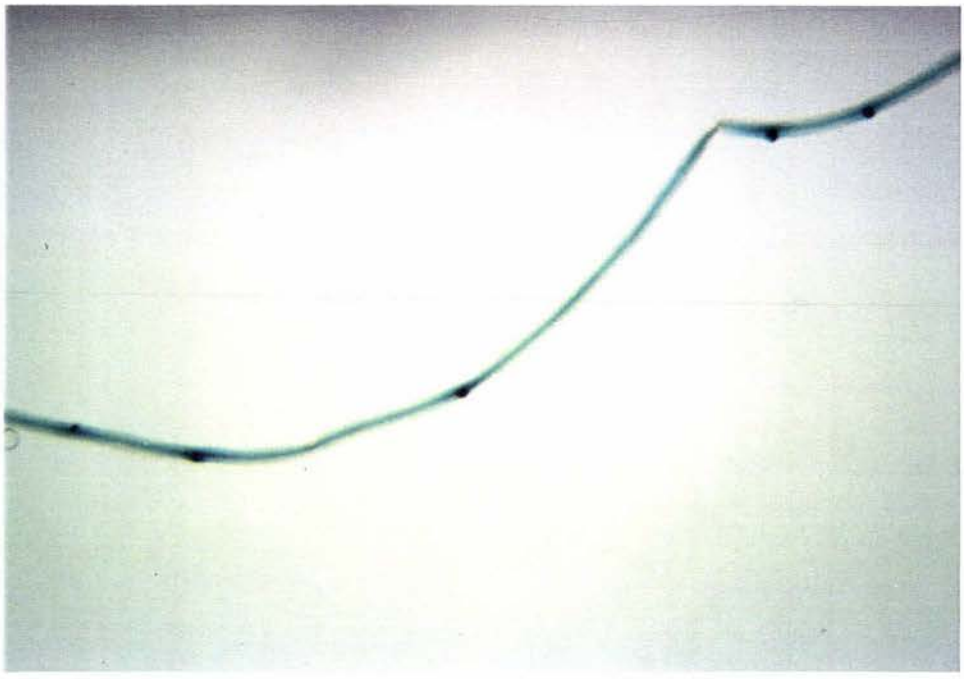


K



L

M



W43.6 was found to be positive in a single root with a dead (brown) root tip, though this plant also exhibited variable GUS activity in floral material (Table 10). Despite the prevalence of shoot apex and root tip staining, little activity was found in flower buds, though the equivalent meristematic activity was probably well before the time of sampling.

While staining was frequently more intense around cut edges of leaf material, this was not necessarily a stain penetration effect as the material was not fixed and a limited metabolic response to wounding could have occurred, until the tissue died during staining (37°C, high sodium concentration).

4.3 Histochemical staining of floral material

Floral material from glasshouse grown plants was also tested for GUS activity. Initially floral material, especially the ovary, underwent oxidative browning during XGluc staining (for example Figure 32 **A**). It was determined that adding 100 mM ascorbate to the staining solution (Martin *et al.*, 1992) reduced browning and that vacuum infiltration was a critical step.

The fixation protocol of Koltunow *et al.* (1990) was compared unaltered (0.1% v/v formaldehyde, 0.1% v/v Triton X-100, 0.1% v/v β-mercaptoethanol, 100 mM phosphate buffer, pH 7.0, Figure 32 **A** and **C**), with 100 mM ascorbate added (Figure 32 **B**) and with ascorbate and phosphate buffer only (Figure 32 **D**). Fixing and vacuum infiltrating for 10 minutes was followed by two washes in 50 mM phosphate buffer (pH 7.0) and then XGluc staining including 100 mM ascorbate. The clearest staining was with ascorbate alone, hence 100 mM ascorbate was chosen as the only fixative agent required.

Subsequently, sections were cut thinner, immersed in phosphate buffer with ascorbate added, before vacuum infiltrating in the final stain solution including 100 mM ascorbate (Section 2.5.2).

The effects of ferricyanide and ferrocyanide were examined. A concentration of

Figure 32. Trial of GUS histochemical detection within floral organs.

Flowers harvested at anthesis from plant A26.6 were cut longitudinally and the ovary/receptacle portion stained with XGluc (Section 2.5.1) with 100 mM ascorbate added after various pretreatments. A single flower was halved between **A** and **B**. Similarly, a flower was halved between **C** and **D**.

A Vacuum infiltration for 10 minutes with 0.1% (v/v) formaldehyde, 0.1% (v/v) Triton X-100, 0.1% (v/v) β-mercaptoethanol, 100 mM phosphate buffer (pH 7.0) followed by two washes in 50 mM phosphate buffer (pH 7.0) and then XGluc staining.

B As in **A** but with 100 mM ascorbate added to the vacuum infiltration treatment.

C Identical pretreatment and staining as **A**.

D Vacuum infiltration for 10 minutes in 100 mM ascorbate, 100 mM phosphate buffer (pH 7.0).

A

B

C

D

A

Figure 33. Effect of ferricyanide/ferrocyanide catalyst on GUS histochemical staining in floral organs.

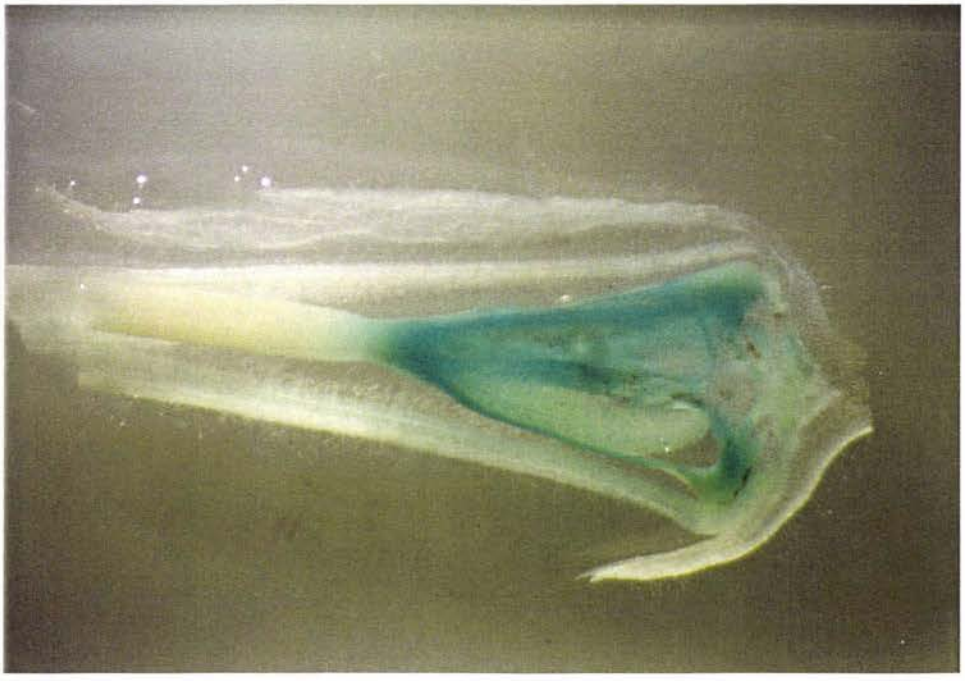
Flowers harvested at anthesis from plant A26.6 were cut longitudinally and the ovary/receptacle portion stained with XGluc (Section 2.5.2) though with variations on the ferricyanide and ferrocyanide concentrations.

A No ferricyanide nor ferrocyanide.

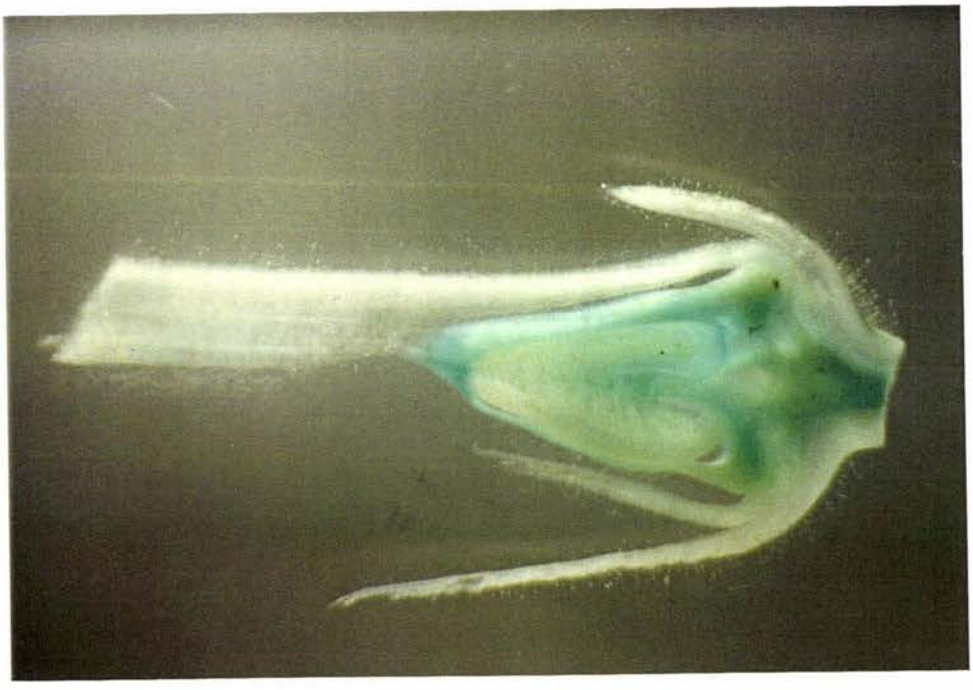
B 0.5 mM ferricyanide and 0.5 mM ferrocyanide included in stain solution.

C 2.0 mM ferricyanide and 2.0 mM ferrocyanide included in stain solution.

B



C



0.5 mM ferricyanide and ferrocyanide (Figure 33 **B**) was found to give better staining than 2.0 mM ferricyanide/ferrocyanide (Figure 33 **C**) while reducing the diffusion of blue precipitate as seen without this catalyst (Figure 33 **A**).

This experiment finalised the staining protocol (Section 2.5.2) for screening floral material. With care, it could give very clean controls (Figure 34). The browning seen in immature anthers (Figure 34 **A**) could be eliminated by sectioning through the anther before vacuum infiltration.

As this protocol differed substantially from that used to screen tissue culture material (Section 2.5.1), the two were directly compared on both floral and tissue culture material, and with further variation utilising 20% (v/v) methanol. Interestingly, optimal conditions were quite different for floral material and tissue culture material. The staining protocol for floral material (with ascorbate and vacuum infiltration) could not be improved with methanol. On the other hand, ascorbate and vacuum infiltration inhibited staining in tissue culture material whilst methanol enhanced it.

Fifty-six plants from the A and W series were histochemically tested for GUS activity in floral organs (Section 2.5.2). Of the twenty-two plants already known to be positive from tissue culture shoot or root material, eleven of these were also positive in floral organs (details below in Section 4.2). From 34 W series plants previously negative in the tissue culture screening, three (9%) were found to be GUS positive in floral tissues. Perhaps not surprisingly, these three had the most defined organ specific expression in floral material; both W11.1 and W48.5 with GUS activity only in immature anthers and W53.1 with GUS activity only in mature style tissue (Figure 35 **A**). While staining patterns varied, the commonest site of GUS activity was the receptacle and base of the ovary (for example W42.1, Figure 35 **B**) and also nectaries as seen with A26.6 (Figure 33). This pattern was not absolute, for example W43.2 did not stain in receptacle tissue, but throughout the surface of the ovary and also in nectaries and style (Figure 35 **C**).

Except for the two plants described with activity in immature anthers, all other staining was confined to the time of anthesis, with no other staining observed in immature flower buds and none at all in maturing seed pods.

A

Figure 34. Trial of XGluc staining of floral organs from untransformed plants.

Floral material from untransformed tobacco was stained with XGluc according to the final staining protocol (Section 2.5.2).

A Longitudinal section through a flower bud

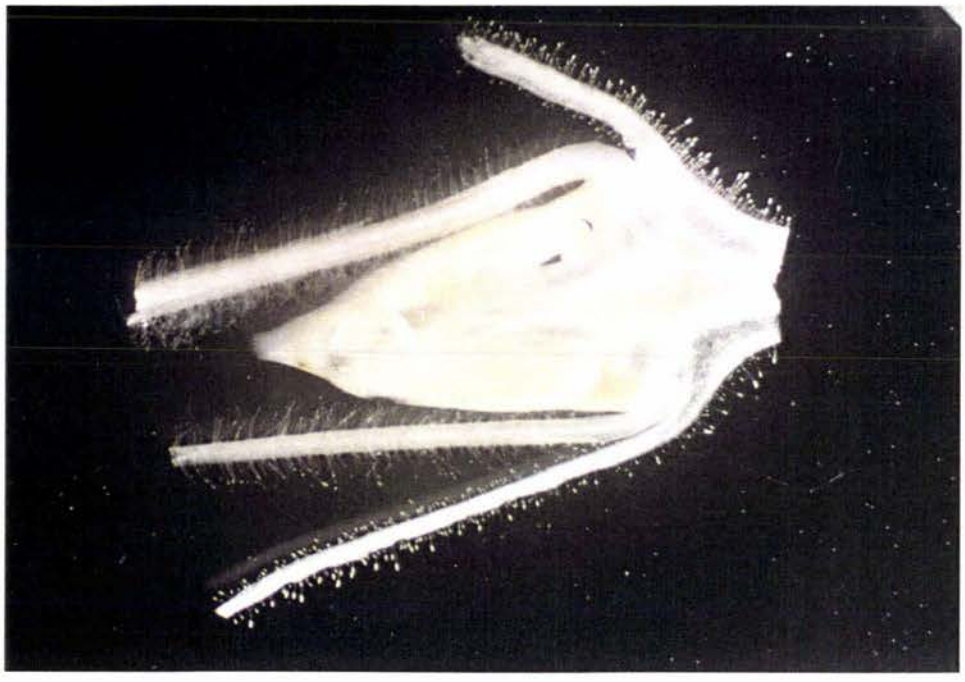
B Longitudinal section through the ovary, receptacle, lower petals and sepals of a flower at anthesis.

C Transverse section through a maturing seed pod.

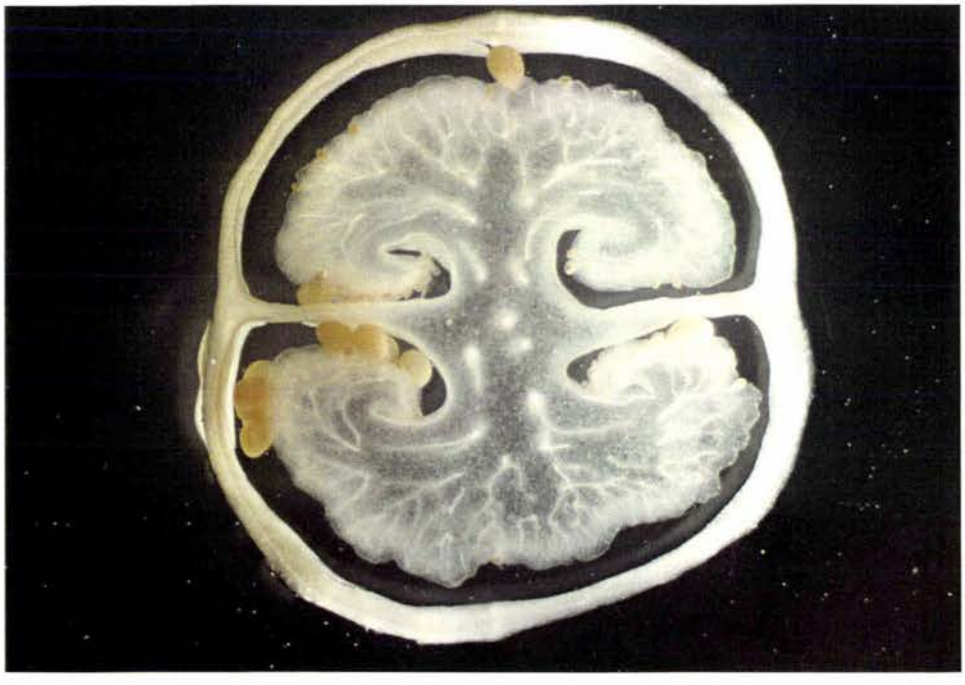
D Stigma and anther from a flower at anthesis.

E Anther and portion of upper petal from a flower at anthesis.

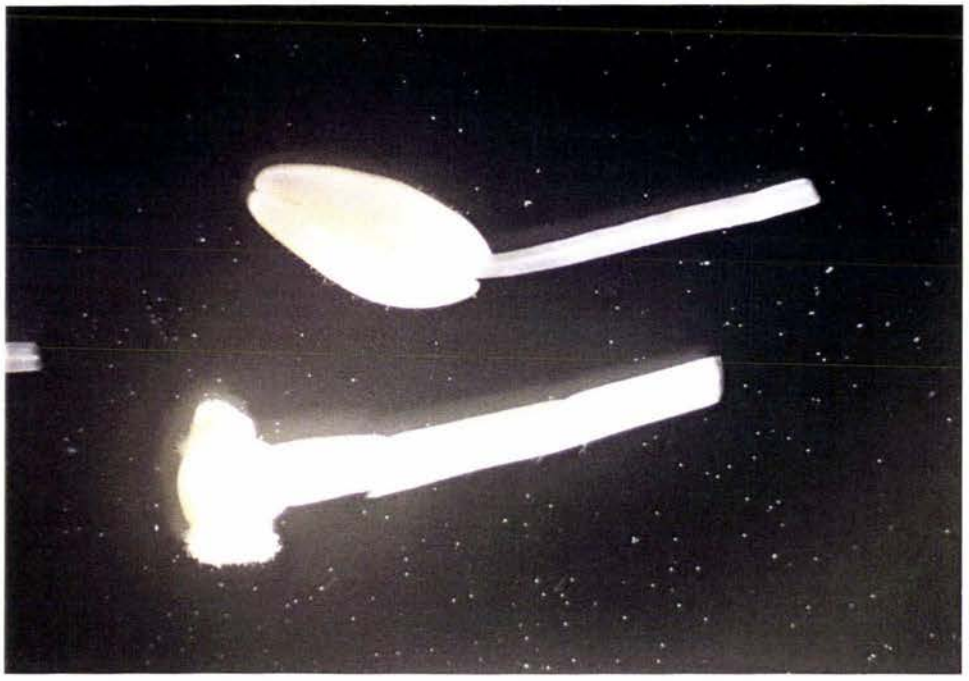
B



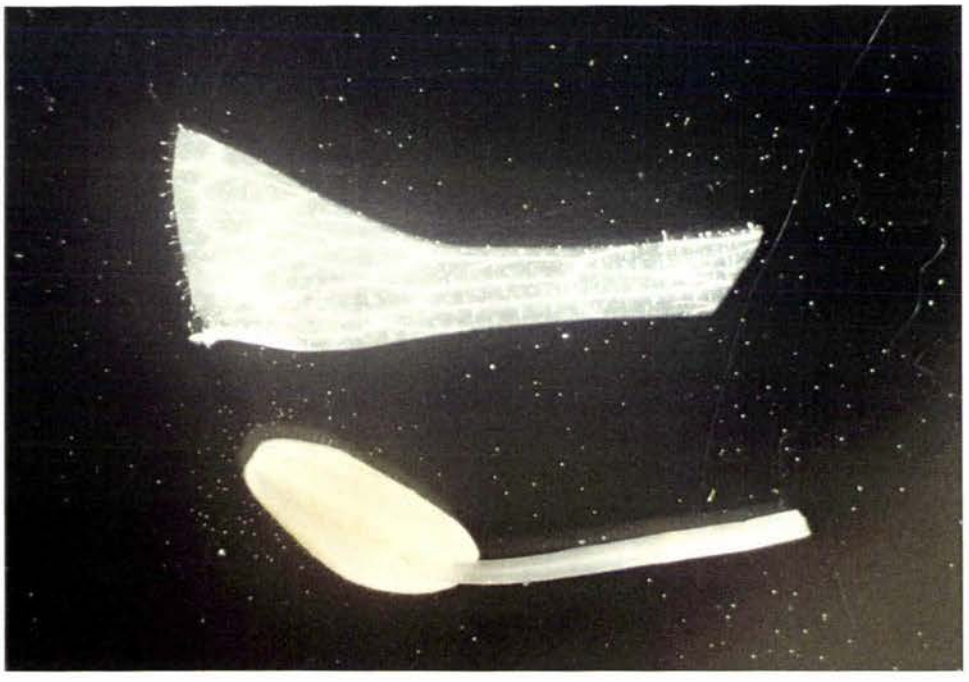
C



D



E



A

Figure 35. GUS activity in floral organs.

Examples of GUS activity detected histochemically (Section 2.5.2) in:

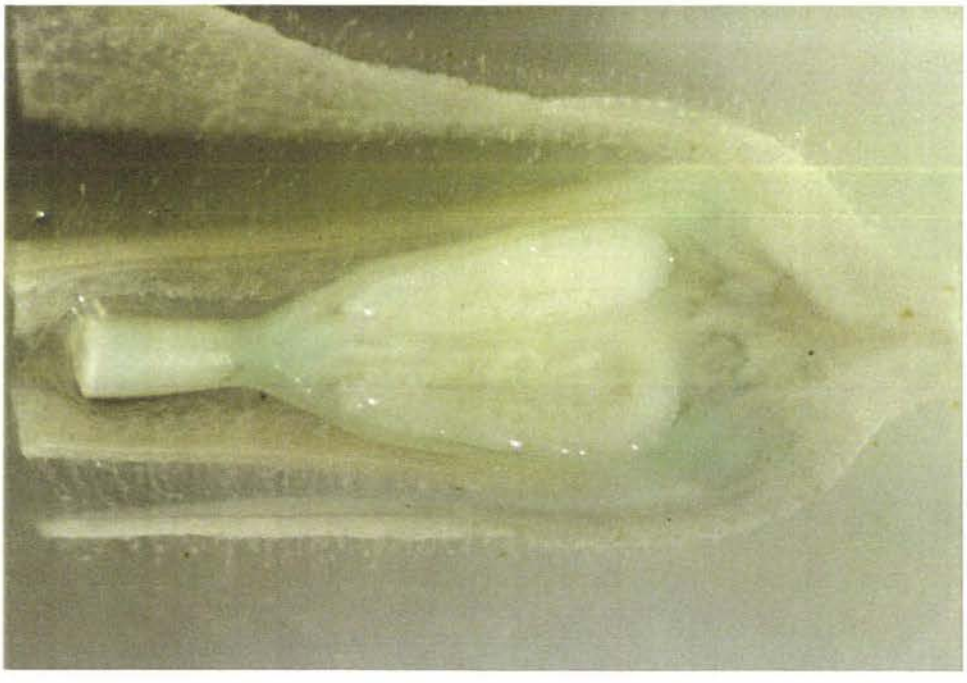
A The upper style of plant W53.1.

B The receptacle and lower ovary of plant W42.1.

C The ovary and nectaries of plant W43.3.

B

C



4.4 Determination of T-DNA copy number by Southern hybridisation

The W and A series of plants were analysed by Southern hybridisation in order to estimate the number of T-DNA copies per transgenic plant. Also, this would allow the identification of any clonal isolates arising from the same initial transformation event.

Plant genomic DNA from A series and W series plants was extracted from tobacco leaf material, restriction digested with *Hin*dIII and *Eco*RI individually, separated on 0.8% agarose gels, blotted onto nylon membranes and then probed with first the *gus* gene, stripped and then hybridised with the *npt*II chimaeric gene (protocols detailed in sections 2.6.1 to 2.6.5).

The *gus* probe was prepared by PCR (Section 2.7.1) with universal and reverse primers from a subclone of pUCK1 containing only the 1.9 kb *Eco*RI fragment of the *gus* coding sequence. The *npt*II probe was prepared by gel extracting (Section 2.3.13) the 1.6 kb *Hin*dIII/*Eco*RI fragment from a digest of pNNPT. This fragment contains the full P*nos*-*npt*II-3′*nos* chimaeric gene.

A full length, and single, copy of the pBin19-GTG T-DNA integrated into tobacco genomic DNA should generate one restriction fragment that is greater than 1.9 kb with an *Eco*RI digest probed with *gus*, one greater than 3.5 kb for a *Hin*dIII digest probed with either *gus* or *npt*II and an internal fragment of 3.2 kb with an *Eco*RI digest probed with *npt*II (Figure 36). This combination of fragments was found in only two of the 44 transgenic plants tested with both probes (Appendix 8.2).

Three of the probe/digest combinations will provide separate indications of T-DNA copy number, two across the right border junction (*gus* probe) and one across the left border junction (*Hin*dIII digest, *npt*II probe). As well, the intensity of the internal band relative to the intensity of other bands gives another indication of multiple copies.

Southern hybridisation of *Eco*RI and *Hin*dIII digested genomic DNA from five plants (Figure 37) revealed a large number of hybridising fragments (Table 11), indicative

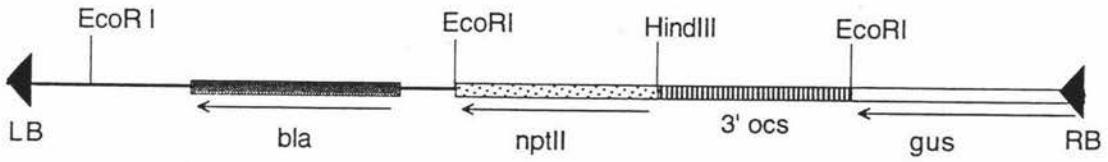


Figure 36. Physical map of the T-DNA of pBin19-GTG.

Abbreviations: LB and RB, left and right borders of T-DNA; *gus*, coding region of β-glucuronidase gene; *npt*II, chimaeric Pnos−*npt*II−3′*nos* gene; 3′ *ocs*, octopine synthase 3′ processing signals; *bla*, β-lactamase gene

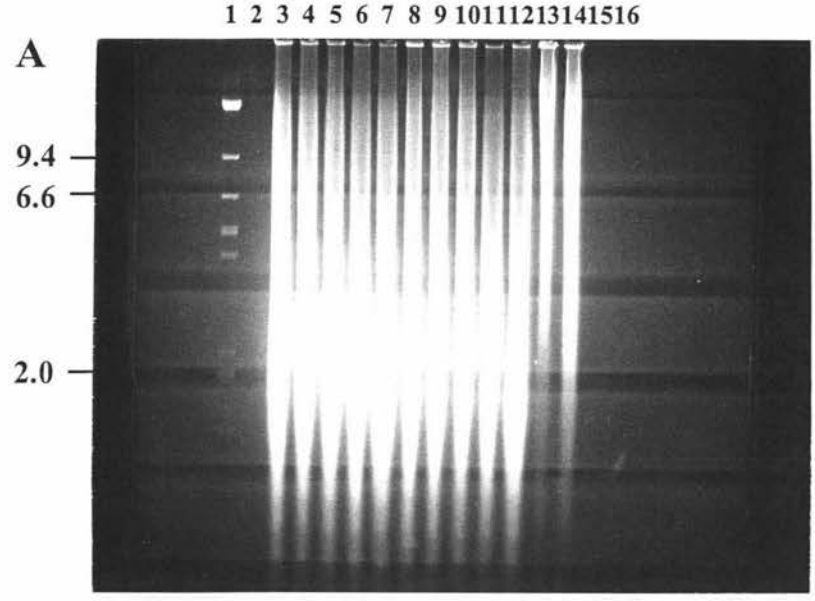


Figure 37. Gel and Southern hybridisation of plant genomic DNA from T-DNA tagged plants.

Plant genomic DNA (approximately 25 μg) was digested with either *Hin*dIII or *Eco*RI, ethanol precipitated and separated by gel electrophoresis on a 0.8% agarose gel. The gel was Southern blotted.

A. Photograph of ethidium bromide stained gel.

B. Autoradiograph of Southern blot probed with *gus* (refer text).

C. Autoradiograph of Southern blot probed with *npt*II (refer text).

Lane 1: *Hin*dIII plus *Hin*dIII/*Eco*RI digested λ DNA size markers (Section 2.3.12).

Lanes 2 and 15: blank.

Lanes 3 to 7 and 13: *Eco*RI digested. Lanes 8 to 12 and 14: *Hin*dIII digested.

Lanes 3 and 8: DNA from plant A1.1.

Lanes 4 and 9: DNA from plant A1.3.

Lanes 5 and 10: DNA from plant A1.4.

Lanes 6 and 11: DNA from plant A18.1.

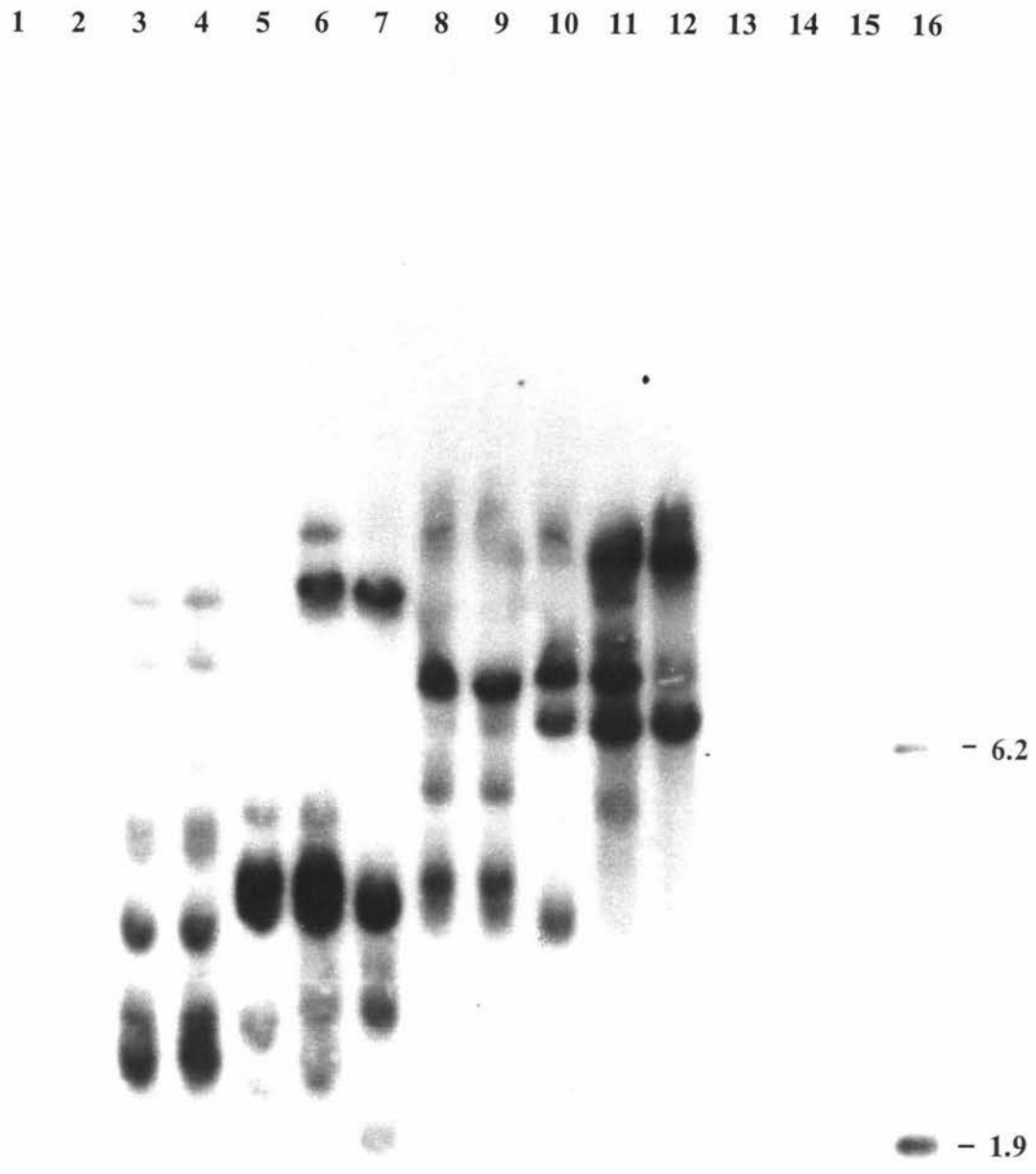
Lanes 7 and 12: DNA from plant A35.1.

Lanes 13 and 14: DNA from an untransformed plant.

Lane 16: Mixture of three digests of pUCK1 DNA; *Hin*dIII, *Eco*RI and *Hin*dIII/*Eco*RI.

Numbers on the right of the autoradiographs are DNA sizes.

B



C

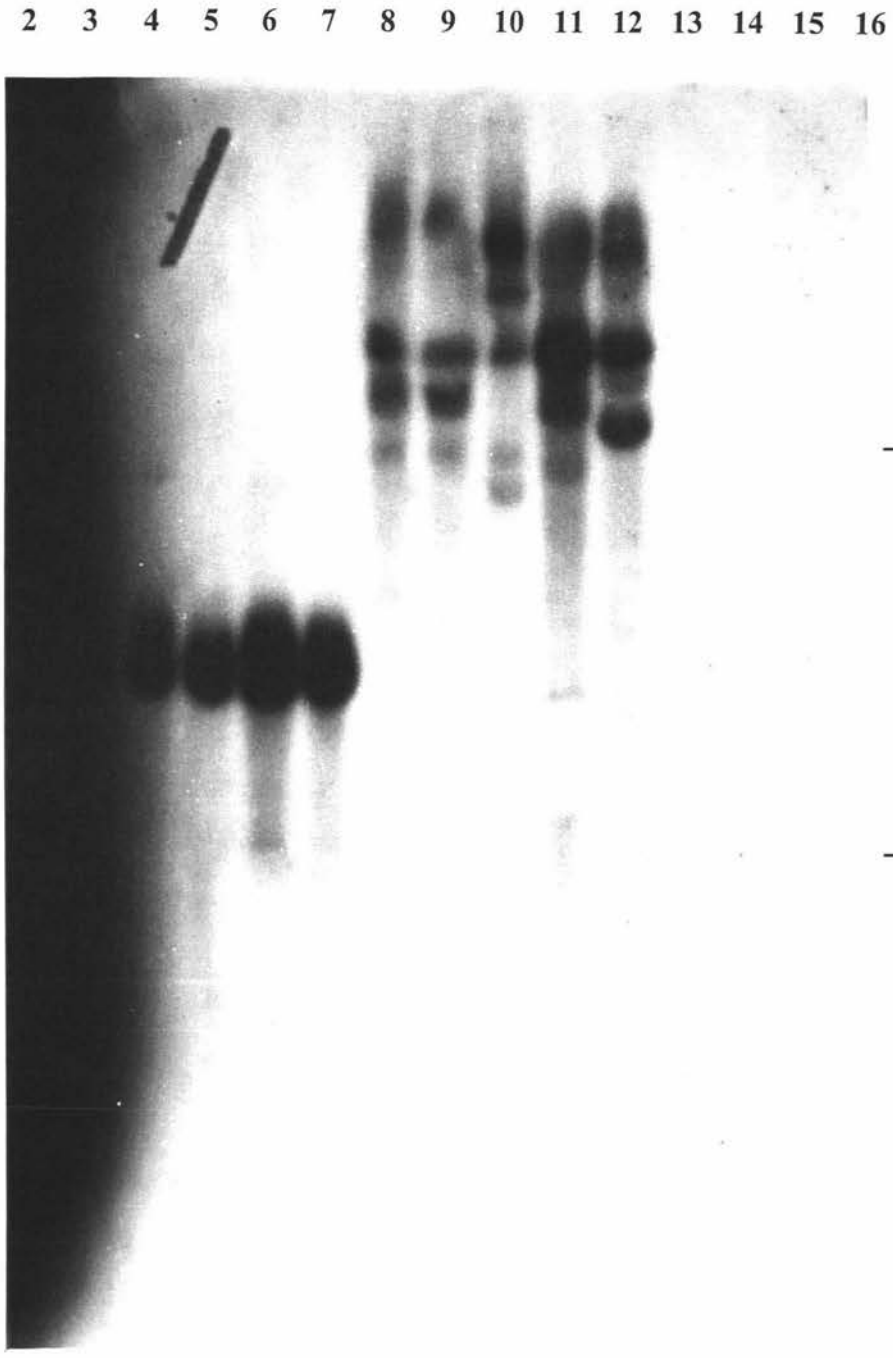


Table 11. Number of significant hybridising fragments from Southern hybridisation of plant genomic DNA from Figure 37.

Probe	Digest	Transgenic plant				
		A1.1	A1.3	A1.4	A18.1	A35.1
gus	*Eco*RI	8	8	4	7	6
gus	*Hin*dIII	7[1]	7[1]	5[1]	8	4[1]
*npt*II	*Eco*RI	1	1	1	1	1
*npt*II	*Hin*dIII	4	4	5[1]	5	5[1]

[1] Possible partial digest on uppermost band in this lane

of multiple T-DNA insertions. *Eco*RI digests probed with *npt*II revealed a strong band of 3.2 kb as expected, with a much weaker band of 6.2 kb (Figure 37 **C** lanes 3 to 7) This is presumably due to the presence of a minor proportion of incompletely digested T-DNA sequences giving rise to an internal T-DNA fragment of 6.2 kb. The weak intensity of this band compared to the 3.2 kb internal band was an indication that hybridising bands seen in *Eco*RI digests with the *gus* probe (Figure 37 **B** lanes 3 to 6) were fully digested products as their relative intensities did not vary by this degree. On the other hand, the intensities were not uniform, which could have three simple explanations. First, partial digestion which is already discounted. Second, internal deletion of part of the hybridising fragment, though the uniformity of internal *Eco*RI band size hybridising to the *npt*II probe in the adjacent section of the T-DNA makes this unlikely. Third, some bands could be represented in more than one copy within the plant genome. This appears most likely, hence there are both multiple T-DNA insertions at different sites in the plant genome and multiple copies (repeats) of particular T-DNA insertions at some sites.

Examination of the *Hin*dIII digests revealed some, especially high molecular weight, fragments hybridising with both *gus* and *npt*II probes (Figure 37, compare between **B** and **C**, lanes 8, 10 and 12). These fragments were discounted as potential partial digests (Table 11) of the internal T-DNA *Hin*dIII site (Figure 36) when they appeared relatively weaker in intensity than the other hybridising fragments. Though it is realised that such fragments apparently hybridising to both probes could also have arisen from tandem repeats of the T-DNA (either inverted about the left border or direct) within the plant genome.

The mixture of pUCK1 fragments (Figure 37 **C** lane 16) did not hybridise with the probe prepared from pNNPT, indicating that the *npt*II probe was not contaminated with pUC8 sequence (the parent vector of both pNNPT and pUCK1) and therefore the probe would not hybridise to the pBR322 derived sequence within the T-DNA. Likewise, the *gus* probe only hybridised to the anticipated pUCK1 fragments (1.9 and 6.2 kb, Figure 37 **B** lane 16) containing the *gus* gene, and not other fragments with only 3′ *ocs* or pUC8 sequence and therefore would not cross hybridise to related sequence within the T-DNA.

After screening all W and A series plants (Appendix 8.2), three pairs of identical plants were found, A1.1 and A1.3 (Figure 37, compare lanes 3 and 4), A88.1 and A88.4, and W5.3 and W5.6 (Appendix 8.2) so only one of each pair has been included in further analysis.

In order to estimate the number of T-DNA insertions containing *gus* sequence, the higher number of either *Eco*RI or *Hin*dIII hybridising fragments was taken, after discounting likely *Hin*dIII partial digests. Including data from both Table 11 and a further 40 plants (Appendix 8.2), the number of *gus* hybridising fragments per transgenic plant ranged from 0 to 8 (Figure 38) with a mean of 3.3. This average is an estimate of right border/*gus* T-DNA fragments. A separate estimate of left border T-DNA ends from *Hin*dIII digested *npt*II hybridising fragments ranged from 1 to 6 per transgenic plant with a mean of 2.7. These two estimates of T-DNA copy number need not agree, as individual T-DNA could have lost internal portions by deletion. As expected, every transgenic plant had at least one *npt*II hybridising fragment which would have been required for the plant to survive selection on kanamycin.

Nine of the 44 plants were A series plants, preselected by screening for GUS activity. This subset had an average of 5.2 right border hybridising fragments against 2.8 for the W series plants alone. This difference was anticipated as plants with a higher frequency of T-DNA insertion are more likely to also have *gus* activation via promoter tagging. Hence GUS positive plants are on average more likely to have larger numbers of T-DNA insertions.

One feature of *Eco*RI digests of genomic DNA from additional plants probed with *npt*II was the range of hybridising bands (when only the internal 1.6 kb fragment was expected), especially lower molecular weight, with some particular fragments of less than 1.6 kb reoccurring among several transgenic plants (Appendix 8.2). This could be attributed to star activity of *Eco*RI, or perhaps caused by a common rearrangement of the T-DNA. However hybridising fragments of less than 1.9 kb were not commonly seen with the same *Eco*RI digests probed with *gus*. Even so, this may have caused a small overestimate of T-DNA copy number.

Additional data from the full set of A and W series plants examined with less than the

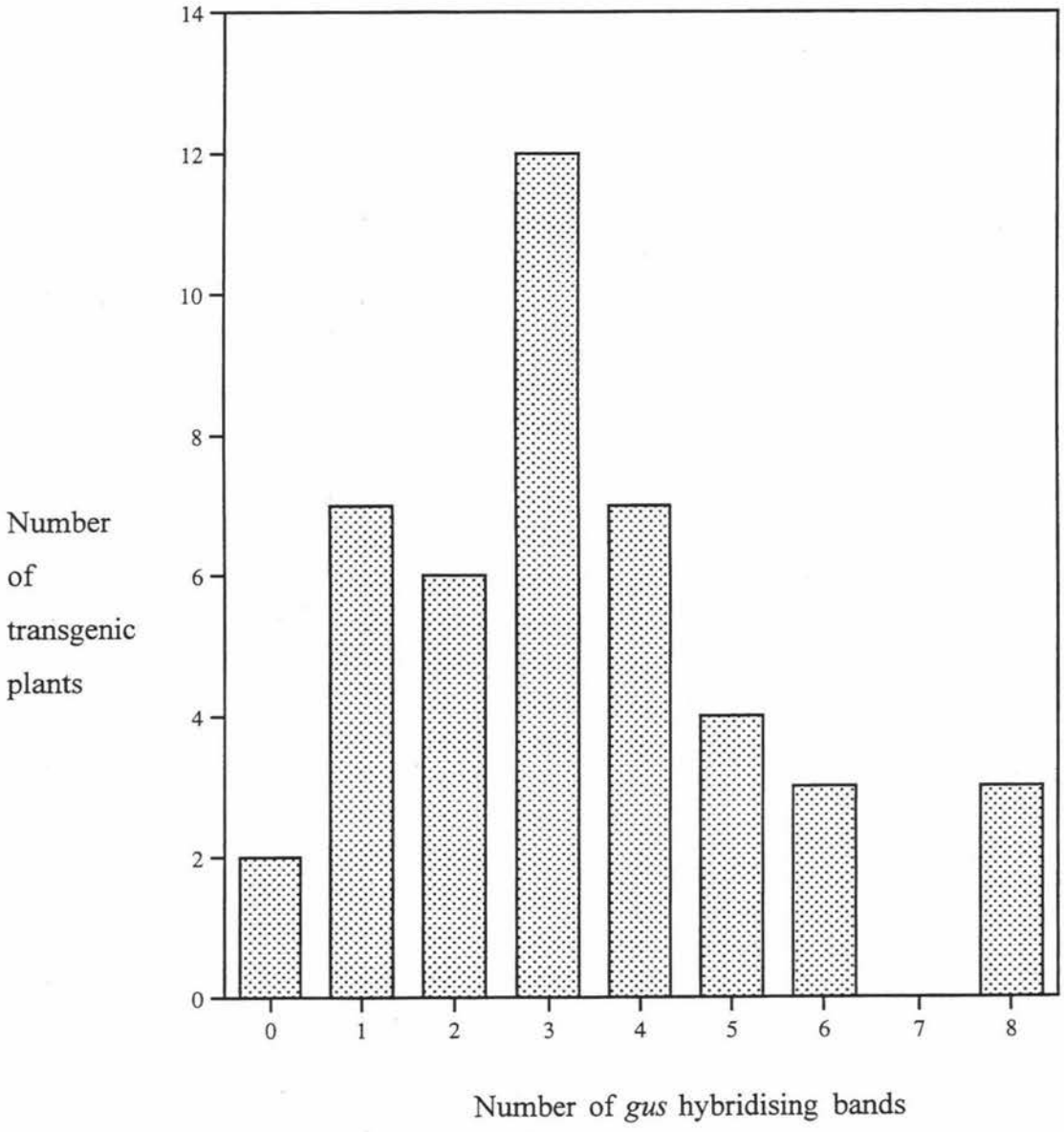


Figure 38. Frequency of *gus* hybridising bands after Southern hybridisation of A and W series transgenic plants.

four probe/digest combinations are appended (Appendix 8.2).

The combination of GUS activity (Sections 4.1 and 4.3) and T-DNA copy (above) now allowed final analysis of active gene tags per T-DNA insertion.

4.5 Overview of promoter tagging information

From 147 plants histochemically screened from the A and W series, 22% were found to have GUS activity in root or shoot material. Another 9% of plants lacking root or shoot GUS activity were found to be GUS positive in floral organs.

The average number of T-DNA insertions from a random sample (W series) was 2.8 per transgenic plant. Hence, on average, 11% of all T-DNA insertions were involved in GUS activation.

5.0 Discussion

5.1 T-DNA promoter tagging vectors

The two new T-DNA promoter tagging vectors, pGTG and pBin19-GTG, constructed in this study (Section 3.1), are comparable to other published plant promoter tagging vectors (Kertbundit *et al.*, 1991; Fobert *et al.*, 1991; Lindsey *et al.*, 1993). All contain a promoterless *gus* reporter gene immediately internal to the T-DNA border and a constitutively expressed selectable antibiotic resistance gene, P*nos*-*npt*II or P35S-*npt*II. These vectors are all suitable for *Agrobacterium*-mediated plant transformation and the generation of a population of transgenic plants. Screening for promoter tagging events revealed GUS activity discernible in a variety of tissues with distinct spatial and temporal patterns (Kertbundit *et al.*, 1991; Fobert *et al.*, 1991; Lindsey *et al.*, 1993) as was found with pGTG and pBin19-GTG (Table 10, Figures 18, 31 and 35).

Although this study provides no direct evidence that this GUS activity was the result of insertions into plant genes, other work (Koncz *et al.*, 1989; summarised by Koncz *et al.*, 1992) has clearly linked T-DNA insertions with gene tagging events, as revealed by cloning of the tagged plant gene. More specifically, GUS activity from a promoter tagging vector similar to pGTG and pBin19-GTG has been linked directly to a T-DNA insertion immediately downstream of a plant gene promoter (Kertbundit *et al.* 1991).

Hence it is extremely likely that the *gus* expression seen after pGTG and pBin19-GTG transformation is driven by promoter sequences found within the plant genomic DNA, rather than an artifact due to T-DNA insertion and re-arrangement causing the *gus* gene to be expressed from regulatory elements incorporated elsewhere within the T-DNA. Within the T-DNA constructions are the promoter of the *nos* gene, the bacterial promoter of *bla*, and, in pGTG only, the promoter of T-DNA gene 5. If any of these promoters were significantly affecting *gus* expression then it might be expected that a repeated pattern of gene expression would result. This was not seen, instead a wide range of patterns of *gus* expression was noted (Table 10, Figures 18, 31 and 35) as have been reported elsewhere (Kertbundit *et al.*, 1991; Lindsey *et al.*, 1993).

Moreover, within the T-DNA regulatory elements, it should be noted that bacterial promoters, excepting T-DNA promoters which are functionally plant promoters anyway, are not known to function in eukaryotes so *bla* can be discounted as a potential contributor to false promoter tagging events. The T-DNA gene 5 promoter was not present in pBin19-GTG and so could not be responsible for the *gus* expression which was observed at comparable frequency with both vectors. The *nos* promoter is regarded as being constitutive (Kuhlemeier *et al.*, 1987). However, even constitutive plant promoters can be made of a variety of functional components which individually display distinct patterns of expression but combined give rise to coverage of a multitude of cell types (Kuhlemeier, 1992). Hence it is conceivable that deletion or duplication of part of the *nos* promoter region and its rearrangement 5′ of the *gus* code could generate a larger range of patterns of gene expression. This level of rearrangement appears extremely unlikely to have both occurred and to have been able to generate the large number of patterns of *gus* expression seen in the transformed plants.

While these promoter tagging vectors give rise to insertion events in which *gus* gene expression is likely to be mediated by plant sequences, there is no guarantee that this expression will reflect the expression pattern or intensity of a normal plant gene promoter. The T-DNA insertion itself may disrupt promoter elements by displacing a 3′ enhancer or a regulatory element within an intron. Additionally, unless the insertion takes place immediately downstream of the TATA box, the *gus* mRNA will be part of the same transcriptional unit as the, now truncated, tagged gene. This transcriptional fusion is likely to lead to a dicistronic message, with the *gus* downstream. Until Koncz *et al.* (1989) demonstrated that such a unit would function it was thought that only one open reading frame and that usually the first, would be translated. Even so, diminishment of expression is still possible.

Another unanticipated development using a promoter tagging vector similar to pGTG and pBin19-GTG has been the discovery of cryptic promoter regions within plant genomic DNA (Fobert *et al.*, 1994). Fobert and colleagues analysed an insertion site which resulted in seed specific expression of *gus*. This region contained no long open reading frame and was not transcribed. Hence at least some *gus* expression may be derived from plant sequences not associated with plant genes.

After modifying pGTG by adding the replication functions of pBin19, the vector was suitable for use in a wider *Agrobacterium* host range than strains carrying pRK replication factors. Therefore strain LBA4404 could be used, which allowed the use of a lower concentration of cefotaxime, potentially avoiding some effects of this antibiotic on plant growth (Valvekens *et al.*, 1988; Borrelli *et al.*, 1992).

The combination of pBin19-GTG in LBA4404 as opposed to pGTG in strain GV3101 also resulted in a higher transformation efficiency. A total of 236 transgenic tobacco plants were screened for GUS activity in both leaf and root material after transformation with one of these promoter tagging vectors, compared to other studies of 171 *Arabidopsis* plants (Kertbundit *et al.*, 1991) and 234 tobacco and 94 *Arabidopsis* plants (Lindsey *et al.*, 1993). However, tobacco transformation is still a time consuming process compared to the *Arabidopsis* seed transformation method of Feldmann and Marks (1987) which is used to generate thousands rather than hundreds of plants. On the other hand protoplast transformation has been used for T-DNA tagging in tobacco (Hayashi *et al.*, 1992) which is suitable for generating tens of thousands of transformants in individual experiments (for example Fritze *et al.*, 1995) and therefore allowing larger scale experiments.

Whilst the promoter tagging vectors in this study are designed for reporter gene/plant promoter fusions across the T-DNA right border, Lindsey *et al.* (1993) chose to place their promoterless *gus* construct immediately inside the T-DNA left border. It is widely reported (for example, Deroles and Gardner 1988b; Bakkeren *et al.*, 1989) that the T-DNA end point is much more variable at the left border. This thesis also found evidence (Section 4.4) that the left border/plant junction was more variable, as a larger number of T-DNA fragments were found that hybridised to sequence near the right border than to a central T-DNA probe. This difference in position of the promoterless reporter gene could cause two differences between their results and the results reported in this thesis. If Lindsey and colleagues had screened plants by Southern hybridisation, they might have observed a higher frequency of T-DNA insertions which contain only their selectable antibiotic resistance gene without the *gus* gene. This would lower the frequency of the number of tagging events seen per transformed plant. On the other hand, a higher percentage of insertion events resulting in *gus* expression would also be linked to their antibiotic marker gene.

Similar to other T-DNA tagging vectors (Forsthoefel *et al.*, 1992; Jiang *et al.*, 1992), especially the pPCV series (Koncz *et al.*, 1989) from which they were derived, pGTG and pBin19-GTG were constructed with an *E. coli* origin of replication and a β-lactamase marker gene for selection on ampicillin that make it possible to clone plant genomic DNA flanking the insertion by plasmid rescue. Alternatively, polymerase chain reaction methods (inverse PCR and single sided ligation mediated PCR) only require suitable restriction sites and synthesis of oligonucleotide primers in order to be able to clone T-DNA flanking sequence.

5.2 Cloning of plant genomic sequence flanking T-DNA insertions

Three techniques, inverse PCR, single sided ligation mediated PCR and plasmid rescue were employed in efforts to clone T-DNA flanking sequences from plants E1 and 118, which were chosen because of their strong root specific GUS activity. These techniques have all been used with success by other workers, although plasmid rescue is the most widely used for cloning after T-DNA tagging (Forsthoefel *et al.*, 1992). It is an especially useful technique in *Arabidopsis thaliana* where the genome is fifty times smaller than in *Nicotiana tabacum* (Lindsey *et al.*, 1993) so that the effective concentration of T-DNA and flanking sequence for transformation into *E. coli* is that much greater.

Inverse PCR was the most useful technique used in this study for the cloning of T-DNA flanking sequences. Three sequences were cloned from plant 118 and two flanking sequences cloned from plant E1 (Section 3.7.1).

The two E1 clones, generated separately from *Eco*RI and *Hin*dIII digests before inverse PCR, around the T-DNA right border, were found to overlap. This fragment corresponds to the smallest of three right border *Eco*RI fragments seen after Southern hybridisation (Section 3.6) and is presumed to be plant DNA flanking one of the multiple T-DNA insertions. No significant open reading frame was present in the sequence, nor was there any significant match against the Genbank DNA database. This sequence could still represent a portion of an unknown plant promoter (with T-DNA insertion upstream of coding sequence) or the sequence driving root specific *gus*

expression is present in one of the two uncloned right border regions.

The 118 flanking sequences represented two T-DNA rearrangements and/or repeats and one unassigned sequence, again presumed to be plant genomic DNA. No significant open reading frame nor homology with the Genbank database was found in this sequence. Again, either the promoter causing *gus* expression in root tissue is to be found in one of the other four uncloned right border regions, or the T-DNA has inserted immediately downstream of the unidentified promoter sequence. Cloning of a large enough fragment followed by screening by plant transformation of this sequence as a fusion with a reporter gene would reveal if it was a functioning promoter.

Inverse PCR appeared to successfully amplify appropriate T-DNA right border sequences, judging from the sequences cloned which included the short fragment of T-DNA between the primer and the right border and from the relatively close match with right border junction fragment sizes observed by Southern hybridisation (Section 3.6).

Whilst DNA impurities were not a problem during PCR, contaminating plasmid DNA was a difficulty, undocumented elsewhere, in plasmid rescue. These experiments were ultimately inconclusive, but it is probable that an internal T-DNA sequence from plant 118, and possibly a left border flanking sequence from plant E1 were cloned (Section 3.7.3). Another limitation of plasmid rescue is the size of the DNA molecule that can be transformed into *E. coli*. With pGTG and pBin19-GTG the usefulness of plasmid rescue was restricted as the origin of replication and ampicillin resistance gene were near the left border, whereas the potential plant promoter region would have been external to the right border sequence, over 8 kb away. The largest plasmid seen after plasmid rescue was 7.5 kb.

In a different approach, Feldmann's insertional mutagenesis experiments utilised T-DNA containing sequence suitable for plasmid rescue adjacent to both the right and left borders (Forsthoefel *et al.*, 1992). This method is not applicable for promoter tagging due to the requirement to place the promoterless marker gene next to one border. Even with this approach, Feldmann (1992) commented that little more than

half of plasmid rescue attempts were successful in cloning a T-DNA segregating with an insertional mutation when the T-DNA was found in concatamers.

Zhixing and Nahon (1995) have recently demonstrated a nine fold improvement in transformation efficiency of large (15 and 20 kb) molecules following DNA gyrase treatment. This refinement, if reproducible with any large DNA molecule (there may be specific gyrase binding sites required), would make further attempts at plasmid rescue of right border flanking sequence worthwhile.

Another potential complication are the reports (summarised by Martineau *et al.*, 1994) that T-DNA insertions into the plant genome could also be flanked by large segments of Ti plasmid. They noted that, in experiments with two different *Agrobacterium*-mediated binary transformation systems, approximately 20% of transgenics contained vector DNA sequence external to the T-DNA borders, as revealed by Southern hybridisation. This would complicate both the identification and cloning of tagged plant genes. They did not comment whether the vector DNA flanked either or both right and left borders. There should be fewer difficulties with promoter tagging than with T-DNA insertional mutagenesis if large sections of vector DNA are regularly transferred flanking the T-DNA as successful promoter tagging ensures that the border sequence around which cloning will focus is already either within the gene or immediately downstream of the promoter.

Messenger RNA anchored PCR (Loh *et al.*, 1989) could have been attempted in order to circumvent the number of T-DNA insertions and instead amplify only the flanking sequence that was present in messenger RNA. This could then be used as a probe against inverse PCR clones to locate the particular flanking sequence involved in *gus* expression.

Another strategy to clone the sequence flanking T-DNA insertion sites would have been to construct genomic libraries and screen these for clones containing T-DNA and especially *gus* sequence. A quicker alternative would be to make use of enhancements to inverse PCR protocols and the increasing length of PCR products that are amplifiable.

Thomas *et al.* (1994) used redundant hexanucleotide restriction endonucleases to generate shorter restriction products over the T-DNA border regions which enabled them to clone more flanking sequences than if they had used standard hexanucleotide restriction enzymes as done in this study. However, incorporating this design into pGTG and pBin19-GTG experiments could require modifications to the plasmids, depending on the availability of enzyme sites. Thomas *et al.* (1994) also utilised nested primers to verify the authenticity of the PCR products before cloning and sequencing. This was not found necessary in this study and indeed, Thomas *et al.* (1994) reported some failures due to T-DNA rearrangements and the loss of a nested primer site. Interestingly, they observed that linearisation was not required for efficient inverse PCR, in contrast to improved yields seen after heat nicking in this study (Section 3.7.1).

Greater PCR product lengths have been obtained, first by modifying reaction conditions and buffers to generate products up to 6 kb (Ponce and Micol, 1992) and later by the addition of a second thermostable polymerase to PCR reactions (Barnes 1994; Westfall *et al.* 1995). This second enzyme contains a proofreading function which is believed to edit the elongating strand and correct misincorporations so that amplification can continue. It is now routine to amplify PCR products over 20 kb in length and hence the largest *Eco*RI bands that hybridised to *gus* in 118 of 10 kb and E1 of 6.3 kb may now be cloned via inverse PCR.

Subsequent to any cloning of potential promoter regions, the clone would need to be verified, by Southern hybridisation to ensure it co-hybridised to a T-DNA right border fragment as well as to a different wild type plant fragment. Then northern hybridisation should ensure that the tissue specificity of the mRNA hybridising to the clone matched that of the observed GUS activity. Finally, subcloning the putative promoter region with a reporter gene and transforming back into a plant should produce the same pattern of gene expression again.

In order to assess whether or not a cryptic promoter (Fobert *et al.*, 1994) had been cloned and is involved in expression of the reporter gene it would be possible, by northern hybridisation, to show both that the cloned sequence is related to a plant gene mRNA in control tobacco plants (with pattern of expression identical to the GUS

reporter gene activity from the tagged plant) and that in the gene tagged plants, an altered mRNA, also hybridising to a *gus* probe, is present. If desired, the coding region of the wild type gene could then be cloned using the section already obtained as a probe against an appropriate library.

Whilst sequences flanking the T-DNA insertion in plants E1 and 118 were cloned, no correlation was made between any of these sequences and the gene promoter activity revealed by *gus* expression in these plants. This was in part due to technical difficulties in obtaining a large enough clone and partly due to the unexpectedly high copy number of T-DNA insertions, such that there were a large number of T-DNA flanking sequences in each plant.

This problem in not new, but was unexpected in this system (refer Section 5.3), being much more of a problem in, for example, mouse gene tagging experiments in which micro-injection leads to multiple copy integrations and structural rearrangements which hamper cloning efforts (Gossler and Zachgo, 1993). At least in plants there is the potential of outcrossing to reduce copy number that is not possible with mouse cell lines. However, outcrossing was not applied to plants E1 and 118 as initially their progeny had no *gus* expression. Later, after 5-azacytidine treatment of 118 progeny was found to permit GUS activity, the parent plants were no longer available.

Due to the importance of the frequency of T-DNA insertion, another population of transgenic plants was generated and assessed for T-DNA copy number.

5.3 T-DNA copy number

An average of 3.3 different right border/*gus* T-DNA fragments were found per transgenic plant, with only 5% of the plants containing a single T-DNA insertion. There were additional indications that some of these insertions were also present in multiple copies.

This was a high copy number of T-DNA insertions, with a very low frequency of single T-DNA insertions, in comparison to other reports. Koncz and Schell (1986)

found the pPCV series vectors (the parent vector of pGTG and pBin19-GTG, refer Section 3.1) gave an average of 1 to 2 copies per plant in *Nicotiana* transformations, with no suggestion of rearrangement or multiple copies from 54 plants tested. Therefore the high frequency of multiple T-DNA insertions and rearrangements was unexpected, especially in light of other transformation results obtained by staff within the same laboratory using identical tobacco genotypes, *Agrobacterium* strain LBA4404 and plasmid vectors based on pBin19. In these experiments, typically one third of transgenic plants contained a single T-DNA insertion (Derek White, personal communication).

One study found a high number of T-DNA copies in *Petunia* (average 2.0 copies per plant, 93 plants screened) though the frequency of single T-DNA insertions (38%) was nevertheless much higher than the results reported here (Deroles and Gardner, 1988a, 1988b). They suggested that the large size of their T-DNA (18 kb) may have contributed to the high level of rearrangement, though my work also found a high frequency of rearrangement with a smaller T-DNA (8.7 kb).

As anticipated (Bakkeren *et al.*, 1989), rearrangements were much more common around the centre/left border portion of the T-DNA molecule, with a lower proportion of T-DNA molecules lacking the right border/*gus* fragment (Section 4.4). This is most dramatically seen with transgenic plants E1 and 118, which contained 3 and 7 (respectively) different T-DNA molecules within plant genomic DNA (Section 3.6) though only 1 and 3 (respectively) independently segregating active kanamycin resistance genes were found (Table 8). Whilst the *npt*II gene could have been silenced, as the *gus* gene apparently was (Sections 3.5.1 and 3.5.2), this result is as easily explained by preferential loss of left border and central portions of the T-DNA. Indeed, partial T-DNA deletion is given as one reason to explain the effective segregating 1.4 kanamycin genes per transgenic plant (Feldmann, 1991) against an average of 3 T-DNA fragments per plant in Feldmann and Marks' (1987) insertional mutagenesis experiment in *Arabidopsis*.

Whilst T-DNA tagging will generate a higher frequency of hits per plant with a higher frequency of insertions, it is regarded as being more useful to obtain plants with single insertions, as cloning and characterisation of any tagged locus is then simplified

(Fobert *et al.*, 1991). Hence the high number of T-DNA insertions obtained with pGTG and pBin19-GTG make them less useful promoter tagging vectors.

In other promoter tagging studies in tobacco, Fobert *et al.* (1991) obtained an average of 1.1 T-DNA copies per transgenic plant (88% single copy) from Southern hybridisation, Lindsay *et al.* (1993) found 33% of plants contained a single segregating kanamycin marker gene, while Koncz *et al.* (1989) determined that 35% of their transformants contained single insertions with their promoter tagging vector from Southern hybridisation. These figures are all significantly higher than the 5% of single copy insertions found in this study, despite the close similarities between the pGTG and pBin19-GTG and the vectors of Koncz *et al.* (1989) which are also pPCV001 derived and the vector used by Lindsey *et al.* (1993) which is pBin19 based.

Fobert *et al.* (1991) used an artificial T-DNA vector in which the right and left borders had been synthesised as isolated 25 base pair sequences. Hence this vector was lacking the right border upstream element known as 'overdrive' (Peralta *et al.*, 1986) which is known to enhance *Agrobacterium* transformation efficiency. I postulated that the presence of the overdrive sequence in pBin19-GTG (and also in pGTG) could be increasing the proportion of multiple T-DNA transformation events. A study was designed by Andrew Griffiths (personal communication) to examine this aspect of *Agrobacterium* transformation. Whilst his results (Andrew Griffiths, personal communication) indicate that removing overdrive sequence decreases the insertion frequency, even with a full overdrive vector, derived directly from pBin19-GTG (only T-DNA sequence modified, left border to right border around the vector sequence is identical), there was a lower frequency (approximately 1.5 T-DNA copies per transgenic plant) of T-DNA insertions.

No simple explanation arises to reconcile the differences between this study and that of Andrew Griffiths which could have caused such a large disparity in T-DNA copy number. Grevelding *et al.* (1993) demonstrated that variation in the plant tissue can affect insertion frequency in *Arabidopsis* where they compared the T-DNA integration frequency between root transformed (11% single copy) and leaf disc transformed material (64% single copy). However no such tissue difference was involved in this case. Indeed, the same tobacco genotype from the same immediate source was used

for transformation, using the same growth conditions and the same vector and T-DNA border sequences. The only major difference was that Andrew Griffiths used a different source of the *Agrobacterium tumefaciens* strain LBA4404, though the LBA4404 isolate used in this study did not give rise to a high insertion frequency in other experiments within the same laboratory (Derek White, personal communication). Other potentially significant variations would include differences in bacterial growth and hence density at inoculation, inoculation time and length of co-cultivation.

A first test to examine the higher copy number of T-DNA insertion with pGTG and pBin19-GTG would be to enlist a co-worker to transform tobacco. Transgenic plants could then be analysed for T-DNA copy to ascertain if the high insertion frequency is peculiar to either these vectors or to tobacco transformation in my hands.

5.4 Frequency of promoter tagging events

This study concurs with the reports (summarised by Koncz, 1992) of a high frequency of promoterless reporter gene activation. Initial screening of leaf and root material, using a microtitre plate fluorometric assay for GUS activity, found 37% of the tobacco plants with *gus* expression. Later, 22% of a new population of transgenic plants were observed to be GUS positive using histochemical screening. Comparing the organ distribution of GUS activity between the two screenings (Sections 3.3 and 4.2) it is apparent that the difference can be attributed to a higher number of weakly GUS positive roots in the fluorometric screening. This raises the question whether the fluorometric screen was more accurate or whether the observed background of higher fluorescence in, especially older, root material led to false positive assignments. For comparison, Lindsey *et al.* (1993), also using tobacco, reported finding *gus* expression in 22% of their leaf samples and 75% of their root samples using a fluorometric assay. While the leaf result is close (19% this study), the frequency of GUS activity in root samples is widely different. Also, in a smaller sample of 24 plants, Lindsey *et al.* (1993) found 92% to be GUS positive in one or more floral organs, once again using a fluorometric assay, compared to 25% from this study's non-random sample of 56 plants already enriched with GUS positive results in leaf and root material, using a histochemical assay. Comparing the individual organ specificities found in flowers

between Lindsey *et al.* (1993) and this study: stamens 66% versus 21%, calyx 21% versus 7%, corolla 8% versus 7% and pistil 42% versus 79%. My sample was partly of plants already known to be GUS positive and this is likely to have heavily skewed my distribution in favour of the pistil, which was only found to be GUS positive in plants also positive in root or leaf material. However there is a difference in stamen figures as well. This tissue was the most difficult to prevent oxidative browning during histochemical staining (for example, Figure 34 **A**); a feature associated with the higher background of methylumbelliferone independent fluorescence seen in older root material.

A high frequency of promoter tagging has also been observed in other plant species including *Arabidopsis thaliana* (Lindsey *et al.*, 1993), potato (Christey *et al.*, 1993; Lindsey *et al.*, 1993) and *Brassica* (Christey *et al.*, 1993).

In this study T-DNA copy was determined by Southern hybridisation (that is, the number of different hybridising fragments ignoring identical sized bands and repeats). Assessing histochemical GUS positive data per T-DNA copy found 11% of T-DNA integrations result in *gus* activation. This figure is lower than the 25 to 30% activation observed by Koncz *et al.* (1989) in both tobacco and in *Arabidopsis* though Fobert *et al.* (1991) found a more similar frequency of about 5% from screening only tobacco leaf material. The difference may in part also be due to the different methods of assessing T-DNA copy number.

The high level of marker gene activation in both *Arabidopsis* and tobacco, despite a fifty fold difference in nuclear DNA content (and comparatively the same number of genes) suggested that the T-DNA is preferentially inserting into transcribed DNA (Koncz *et al.*, 1989; Kertbundit *et al.*, 1991; Lindsey *et al.*, 1993). It was postulated (Van Lijsebettens *et al.*, 1991; Forsthoefel *et al.*, 1992) that different genes might be available for tagging depending on the transformation system used, for example between the *Arabidopsis* seed transformation system and tissue culture transformation, as different sets of genes would be transcriptionally active at the time of transformation. There has been no evidence that this occurs, though different mutagens (T-DNA versus chemical) give rise to different frequencies of mutation in some loci (Koncz *et al.*, 1992). Instead, evidence is accumulating that transposable

elements, including the T-DNA, preferentially insert into certain genomic regions, distinguished on the basis of their specific base content (Capel *et al.*, 1993; Fobert *et al.*, 1994), that is adenine and thymine rich. These regions also contain the plant genes. It remains to be determined why transposable elements preferentially insert into such regions, but the current model suggests that these regions have a less condensed chromatin structure, also associated with their role as transcribed regions, which makes the DNA more available for recombination with the transposable element (Capel *et al.*, 1993). If true, this makes likely the possibility of targeting most genes via gene tagging, excepting those found in other regions such as in heterochromatin.

5.5 Patterns of tagged gene expression

The most significant feature in common with other studies (Fobert *et al.*, 1991; Kertbundit *et al.*, 1991; Lindsey *et al.*, 1993) of the patterns of promoter tagged gene expression is the diversity observed. This sort of variety of patterns has been seen in other molecular studies, such as that of Koltunow *et al.* (1990) following anther specific gene expression using cDNA probes.

While Fobert *et al.* (1991) analysed *gus* expression throughout a small number of seedlings, other studies have not assessed GUS activity specifically within the shoot apex. As could be expected, a range of combinations of promoters expressing both in the shoot apex and in other organs were observed (Table 10). These included *gus* expression within the shoot apex and young leaves, the apex and throughout leaves, and the apex and throughout root, leaf and some floral organs, as well as apex expression alone. While apex expression was in one plant also correlated with root tip expression (W26.1) and in another combined with lateral root initial expression and weak activity in other organs (A41.1), no plant was found in this sample with expression only in the root and shoot meristems at the root tip, lateral primordia and in the shoot apex.

Comparing the patterns of root gene expression observed in this study (Table 10 and Figures 18 and 31) to examples in the literature (Keller and Lamb, 1989; Taylor and Scheuring, 1994; Smith and Fedoroff, 1995; van den Berg *et al.*, 1995) reveals some

similarities and some new patterns of expression. Expression in the roots of plant A41.1 (Table 10) was similar to that of the *RSI-1* gene (Taylor and Scheuring, 1994), being present in lateral root primordia and more weakly in vascular tissue. However plant A41.1 also expressed in shoot material. Plant E1 transiently expressed *gus* in lateral root initials (Figure 18) in a pattern comparable to the hydroxyproline-rich glycoprotein expression described by Keller and Lamb (1989). Whilst root tip expression was often accompanied with expression in lateral primordia, several exceptions were found (plants W42.1, A10.2 and A88.1; Table 10). *Gus* expression usually did not persist at the branch point after the lateral root had grown out, but plant 118 is an exception. Staining for GUS activity remained intense after the lateral root had grown out (Figure 18 **A**). This pattern is, to my knowledge, undescribed elsewhere.

Another previously unreported observation is the variability of *gus* expression seen in some of the promoter tagged plants. Wide variation, especially silencing, is frequently observed with many foreign genes transformed into plants (Finnegan and McElroy, 1994), but in a promoter tagging system it still needs to be determined what proportion of variability is due to the tagged promoters normal function or to a silencing effect. It seems likely that the native expression patterns of at least some tagged promoters will not be observed due to silencing.

It will now be possible to observe finer detail of *gus* expression by histochemical assay followed by microtome sectioning and also to follow gene regulation during development including in progeny. As well, it is likely that further screening for GUS activity under specific environmental stimuli (auxin application, infection etc) or in specific developmental stages would find more tagged gene promoters. This was demonstrated in the floral screen which found three plants with GUS activation not previously seen during screening of somatic organs. Topping *et al.* (1994) have shown the effectiveness of this approach to identify and analyse promoters specific to embryogenesis in *Arabidopsis*. Goddijn *et al.* (1993) used promoter tagged genes to follow modulation of gene expression in nematode induced root structures in *Arabidopsis*.

Particular patterns of tagged gene expression can themselves be used as markers of

cell identity. Van den Berg *et al.* (1995) employed an *Arabidopsis* line containing vascular specific *gus* expression, created through gene tagging with an enhancer trap construct, as a marker. Staining for the tagged gene expression after laser ablation of individual cells allowed the differentiation patterns to be traced.

5.6 Promoter tagging compared with T-DNA mutagenesis

An advantage of promoter tagging over gene tagging via insertional mutagenesis is that promoter tagging provides information on the spatial and temporal pattern of gene expression. Both techniques generate insertional mutagens. However, *Nicotiana tabacum*, used in this study, is amphidiploid, arising from natural hybridisation of *N. sylvestris* and *N. tomentosiformis* ancestors (Okamuro and Goldberg, 1985) and the two parental genomes have not undergone any extensive recombination (Matassi *et al.*, 1991). Because most genes are represented twice in the genome, self-fertilisation will not produce a recessive mutant after insertion into one locus.

When screening for insertional mutations it may be difficult to distinguish all new phenotypes (Feldmann, 1991; Van Lijsebettens *et al.*, 1991) and there is a possibility of false positives that must be excluded by a second round of screening. The mutant must have an observable phenotype to be found in this manner, as opposed to promoter tagging with a *gus* marker gene, when any tagged gene expressed at the time of histochemical screening can be discovered. It remains to be determined whether genes, perhaps involved in gene regulation as transcription factors and transcribed at low levels, can be distinguished by histochemical screening, or whether these genes will only be found by insertional mutagenesis. In the latter case, recessive mutations would cause a significant mutation to be more easily observed. In contrast, it is possible a gene could play a vital developmental role but is functionally redundant within a gene family, as has been demonstrated after insertional inactivation of certain genes involved in mouse embryogenesis (Copp, 1995). Insertional mutagenesis would not be able to identify such genes, though promoter tagging could.

The most significant advantage gained from insertional mutagenesis is that the mutant phenotype gives some clue as to the physiological or genetic role of the tagged gene,

whereas promoter tagging reveals its pattern of expression.

Two significant advantages of promoter tagging are that the screening takes place in the primary transgenic, avoiding another plant generation, and that false positives due to other mutational events are excluded from the process as only promoters activating *gus* expression will be selected. Considering the high frequency of mutations unlinked to T-DNA insertions (in the range of 50 to 90%; Forsthoefel *et al.*, 1992; Koncz *et al.*, 1992) this could lead to many false leads in an insertional mutagenesis programme.

It should be considered that successful promoter tagging requires insertion of the T-DNA in the appropriate orientation downstream of the TATA box, whereas insertional mutagenesis has a potentially much larger target region, in either orientation and including the promoter region. Fobert *et al.* (1994) reported that the non-coding region from -200 to -1 upstream of the TATA box is particularly adenine/thymine rich compared to the coding region from +1 to +200 in known flowering plant genes. This AT rich region may be a more suitable target for T-DNA integration as discussed above. The actual integration site of more tagged genes needs to be examined before a conclusion can be reached.

The two techniques of promoter tagging and insertional mutagenesis complement each other and allow variations on the theme of characterising and isolating plant genes. This is demonstrated in a report on the cloning of a plant gene using another variation of gene tagging. Springer *et al.* (1995), using exon trapping (little different from promoter tagging in this regard) in *Arabidopsis* were able to demonstrate that a tagged gene, while leading to megagametophyte death when present in two copies, was also expressed throughout the life cycle of the plant.

Another consideration raised by Fobert *et al.* (1994) is that promoter tagging with a transcriptional fusion vector, as used in this study, will lead to cloning of an as yet unknown proportion of cryptic promoters. This will be avoided in mutagenesis studies as the screened mutant phenotype must be the result of insertion into a genuine gene. Alternatively, translational fusions, or exon trapping vectors should be used with most of the benefits of promoter tagging, including screening in the primary transgenic, but potentially with a lower frequency of hits. Kertbundit *et al.* (1991) found activation

frequencies of 2% and 54% for translational and transcriptional fusion tagging respectively, which supports the hypothesis that there are a large number of cryptic promoters. However Koncz *et al.* (1992) reported much more similar frequencies of 32% and 31% for their translational and transcriptional fusion vectors.

Hence the frequency of tagging either cryptic promoters or plant genes needs to be resolved, as well as the location of T-DNA insertions in plant genes, to determine whether promoter tagging has the same target range as insertional mutagenesis.

5.7 Silencing and methylation

After observing the absence of *gus* expression in progeny of E1 and 118, it was decided to ascertain the effect of 5-azacytidine on *gus* expression, due to the increasing body of literature pointing to an involvement of methylation in gene silencing (for example: Hobbs *et al.*, 1990; Jorgensen, 1990; Weber *et al.*, 1990; Matzke and Matzke, 1991; Bochardt *et al.*, 1992; Meyer *et al.*, 1992).

Gene expression can be altered with 5-azacytidine, which is thought to act by inhibiting methyltransferases, but it should be noted that the compound is cytotoxic, can alter chromosome morphology and affects differentiation in an organism that lacks 5-methylcytosine (reviewed by Jones, 1985).

Plant DNA may have as many as 32% of cytosine residues methylated, compared to about 8% in animals (Hepburn *et al.*, 1987). In plants, the methylation sites include both CG dinucleotides and CNG trinucleotides, which allow faithful transmission through symmetry to daughter chromatids (Hepburn *et al.*, 1987), and non-symmetrical cytosine residues (Meyer *et al.*, 1994). In some circumstances 5-methyl adenine has also been found in plant genomic DNA (Muller *et al.*, 1990).

In the experiments reported in this thesis, 5-azacytidine treatment temporarily restored *gus* expression in progeny of plant 118 but not plant E1. Plant 118 has seven different right border/*gus* fragments indicating that there are at least seven different T-DNA sequences incorporated into the plant genomic DNA.

This situation appears analogous to that of Hobbs *et al.* (1993) who demonstrated the presence of two types of T-DNA insertions, those with an additive effect on *gus* gene expression, and those with an over-riding silencing of gene expression. The silencing insertions contained full length *gus* genes, but with double T-DNA copies at the one locus. This study has not demonstrated the presence or absence of this form of T-DNA insertion, but with the number of T-DNA copies present it is a possibility. Progeny of plant E4 segregated for GUS activity in a Mendelian manner, with a 1:2:1 ratio of high:low:absent GUS activity. This is typical of an additive effect of *gus* expression from a single copy of *gus* (Hobbs *et al.*, 1993).

Increased methylation has also been correlated with gene silencing (Hobbs *et al*, 1993). Other studies (Weber *et al.*, 1990; Matzke and Matzke, 1991; Bochardt *et al.*, 1992; Meyer *et al.*, 1992) have shown that methylation is directly involved in gene silencing, as supported by the observation in this thesis that 5-azacytidine treatment restored gene expression.

Plant E1 did not respond to 5-azacytidine treatment, but this may be due to a more limited expression pattern of the controlling plant promoter element, which expressed only in lateral branches in the primary transgenic. Hence *gus* expression may never have been seen in the radicle. Further, the window for 5-azacytidine influence appeared very limited, in that GUS activity was restored at three weeks post germination in some 118 seedlings, but this activity was lost again by six weeks. It would be possible to further analyse E1 progeny by growing the plants in liquid culture to maintain continuous 5-azacytidine treatment.

In addition, to demonstrate that the effect of 5-azacytidine is through methylation, it would be possible to outcross the progeny to reduce copy number and examine the individual genes via genomic sequencing (Church and Gilbert, 1984; Clark *et al.*, 1995) in order to correlate the presence or absence of methylation with *gus* expression.

It appears that methylation is a defence mechanism for inactivating foreign DNA within a wide range of eukaryotic cells (Meyer and Heidmann, 1994; Bird, 1995). Other mechanisms for gene silencing are also known or under study (reviewed by Finnegan and McElroy, 1994) including post-transcriptional control of messenger

RNA processing.

Another problem of plant transformation indirectly addressed by this study is the so-called position effect (Dean *et al.*, 1988), in which independently transformed plants exhibit different levels of transgene expression, or silencing. It appears that the high frequency of T-DNA insertion into plant genes, as shown here by the high frequency of promoter tagging per insertion (11%), will be placing the introduced genes in the immediate vicinity of plant regulatory elements. These elements could be disrupting independent expression of the introduced gene. As well, the effective silencing of some of these introduced T-DNA elements by methylation, as demonstrated here by the re-activation of the gene activity with 5-azacytidine, or silencing through another mechanism (Finnegan and McElroy, 1994), could be leading to variation in transgene expression between individual transformants. Indeed, there is no reason that these mechanisms couldn't interact, for example methylation, chromatin structure and transcriptional availability (Selker, 1990).

6.0 Summary and conclusions

The plasmid pPCV604 (Koncz, 1989), obtained from Dr Csaba Koncz, was used as the starting point for a new series of promoter tagging vectors. A partial digest of pPCV604 removed a short fragment and allowed the kanamycin resistance gene from pBin6 (Bevan, 1984) to be added, creating pGT. Then the *gus* coding sequence with 3′ octopine synthase termination signals from pKIWI101a (Janssen and Gardner, 1989) was ligated into pGT, replacing the promoterless hygromycin phosphotransferase gene with a promoterless *gus* gene, creating pGTG. This binary transformation vector does not contain full replication functions and is maintained with pRK replication genes on either a helper plasmid or chromosomally integrated. In order to bypass this restriction, the T-DNA of pGTG was combined with the vector sequence of pBin19 (Bevan, 1984) creating pBin19-GTG. This plasmid contains its own replication genes derived from pBin19.

Both pGTG and pBin19-GTG promoter tagging vectors have a promoterless *gus* gene with initiation codon 62 base pairs inside the T-DNA. The sequence between the right border element and the initiation codon includes translation termination codons in all three reading frames. Therefore, insertion of the T-DNA into a plant gene can only lead to activation of the *gus* gene via a transcriptional fusion with a plant gene.

The plasmid pBin19-GTG in strain LBA4404 was found to have a higher *Agrobacterium tumefaciens*-mediated *Nicotiana tabacum* transformation efficiency than pGTG in replication permissive strain GV3101, making the former combination more useful for tobacco transformation.

Nicotiana tabacum leaf segments were transformed with pGTG or pBin19-GTG and transgenic plants were selected for resistance to kanamycin. A population of 87 transgenic tobacco plants were fluorometrically screened using a microtitre plate assay for GUS activity in leaf and root material and 37% were found to contain GUS activity, indicating a high frequency of promoter tagging.

Two transgenic plants, E1 and 118, with root specific *gus* expression were further analysed. *Gus* expression, tested histochemically, in both plants was found in root tips

and lateral root primordia, but the intensity and details differed. In plant 118, GUS activity was found at lateral root branch points from the main root axis, but not in E1. Progeny after self-fertilisation lacked GUS activity, though this could be restored in progeny of 118 with 5-azacytidine treatment, suggesting methylation was involved in the gene silencing. Restoration was only temporarily seen at three weeks after imbibition and was absent again at six weeks.

Southern hybridisation, inverse PCR cloning of T-DNA flanking sequences and segregation on kanamycin indicated the presence of multiple T-DNA copies within most of the primary transformants. Furthermore, inverse PCR sequence from one plant indicated multiple and truncated T-DNA insertions at one or more loci.

A further population of transformed plants was generated with pBin19-GTG and histochemically screened for GUS activity. Among 147 plants tested, 19 were found to be positive in shoots only, six positive in roots only and eight positive in both shoots and roots. Staining of floral material was optimised by vacuum infiltration of stain solution including 100 mM ascorbate. Fifty-six plants were histochemically tested in floral organs. Three plants previously negative in shoots and roots were found to be positive in floral organs. Another 11 transformants, previously positive from shoot or root testing, were also positive in floral material. A diverse range of patterns of *gus* expression were observed and described, including new patterns involving root branching.

Overall, an average of 33% of plants were found with GUS activity in one or more organs. This frequency of gene activation is within the range previously reported in other promoter tagging studies in tobacco once allowance has been made for T-DNA copy number and organs sampled (5%, leaf only, Fobert *et al.*, 1991; 30%, leaf and root, Koncz *et al.*, 1992) with the exception of Lindsey *et al.* (1993) who found high frequencies (75%, roots; 92%, flowers) to be GUS positive using a fluorometric screen. However their GUS positive leaf frequency (22%) was similar to that observed in this study (19%).

Forty four plants from this population were analysed for T-DNA copy via Southern hybridisation with a *gus* probe (right border junction T-DNA) and *npt*II probe (central

T-DNA). Multiple copies were frequently found with an average of 3.3 T-DNA copies per transgenic plant. Overall, an average of 11% of T-DNA copies were found to be involved in *gus* activation.

The difference between the fluorometric (37% GUS positive) and histochemical (22% GUS positive) screens for GUS activity in root and shoot material was discussed. The difference was possibly due to false positive GUS determinations in root material in the fluorometric screen, which had found a higher number of weak GUS positive roots. It is suggested that care is needed in assigning promoter tagging hits from fluorometric screening.

Variable expression was observed with promoter tagged genes. It is suggested that further research is required to determine whether this variation was due to silencing, perhaps by methylation, or was a result of the normal expression patterns of the tagged promoters.

It is also suggested that additional research is required to determine the proportion of cryptic promoters (Fobert *et al.*, 1994)

7.0 References

Advisory committee on novel genetic techniques (1982) Revised New Zealand guidelines for genetic manipulation research. Ministry of Science, Wellington. 28p.

Advisory committee on novel genetic techniques (1994) New Zealand code of practice for small scale genetic manipulation research. Ministry for the Environment, Wellington. 40p.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology **215** 403-410.

Andre D, Colau D, Schell J, Van Montagu M, Hernalsteens J-P (1986) Gene tagging in plants by a T-DNA insertion mutagen that generates APH(3')II-plant fusions. Molecular and General Genetics **204** 512-518.

Baker B, Schell J, Lorz H, Fedoroff N (1986) Transcription of the maize controlling element "Activator" in tobacco. Proceedings of the National Academy of Sciences, USA **83** 4844-4848.

Bakkeren G, Koukolikova-Nicola Z, Grimsley N, Hohn B (1989) Recovery of *Agrobacterium tumefaciens* T-DNA molecules from whole plants early after transfer. Cell **57** 847-457.

Barnes WM (1994) PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. Proceedings of the National Academy of Sciences, USA **91** 2216-2220.

Beringer JE (1974) R factor transfer in *Rhizobium leguminosarum.* Journal of General Microbiology **84** 188-198.

Bevan M (1984) Binary *Agrobacterium* vectors for plant transformation. Nucleic Acids Research **12** 8711-8721.

Bingham PM, Levis R, Rubin GM (1981) Cloning of DNA sequences from the *white* locus of *D. melanogaster* by a novel and general method. Cell **25** 693-704.

Biochemical Journal (1992) Policy of the journal and instructions to authors for 1992. Biochemical Journal **281** 1-19.

Biorad Laboratories (1989) Bacterial electro-transformation and pulse controller instruction manual. Biorad Laboratories, Richmond, California. 34p.

Bird AP (1995) Gene number, noise reduction and biological complexity. Trends in Genetics **11** 94-100.

Bochardt A, Hodal L, Palmgren G, Mattsson O, Okkels FT (1992) DNA methylation is involved in maintenance of an unusual expression pattern of an introduced gene. Plant Physiology **99** 409-414.

Bolivar F, Rodrigues RL, Greene PJ, Betlach MC, Heyneker HL, Boyer HW, Crosa JH, Falkow S (1977) Construction and characterisation of new cloning vehicles. II. A multi-purpose cloning system Gene **2** 95-113.

Borrelli GM, Di Fonzo N, Lupotto E (1992) Effect of cefotaxime on callus culture and plant regeneration in durum wheat. Journal of Plant Physiology **140** 372-374.

Boyer HW, Roulland-Dussoix D (1969) A complementation analysis of the restriction and modification of DNA in *Escherichia coli.* Journal of Molecular Biology **41** 459-472.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principles of protein-dye binding. Analytical Biochemistry **72** 248-254.

Capel J, Montero LM, Martinez-Zapater JM, Salinas J (1993) Non-random distribution of transposable elements in the nuclear genome of plants. Nucleic Acids Research **21** 2369-2373.

Chilton M-D, Currier T, Farrand SK, Bendich AJ, Gordon MP, Nester EW, (1974) *Agrobacterium tumefaciens* DNA and PS8 bacteriophage DNA not detected in crown gall tumors. Proceedings of the National Academy of Sciences, USA **71** 3672-3676.

Christey MC, Conner AJ, Tynan JL, Williams MK, Sinclair BK, Shum TS, Clark P (1993) Frequency of transcriptional gene fusion during *Agrobacterium*-mediated transformation of potato and *Brassica.* International Association for Plant Tissue Culture, New Zealand Branch, Conference Presentation. February 17-19, Levin, New Zealand.

Chuck G, Robbins T, Nijjar C, Ralston D, Courtney-Gutterson N, Dooner H (1993) Tagging and cloning of a petunia flower color gene with the maize transposable element *activator.* The Plant Cell **5** 371-378.

Church GM, Gilbert W (1984) Genomic sequencing. Proceedings of the National Academy of Sciences, USA **81** 1991-1995.

Clark SJ, Harrison J, Frommer M (1995) CpNpG methylation in mammalian cells. Nature Genetics **10** 20-27.

Cooley L, Kelley R, Spradling A (1988) Insertional mutagenesis of the *Drosophila* genome with single P elements. Science **239** 1121-1128.

Copp AJ (1995) Death before birth: Clues from gene knockouts and mutations. Trends in Genetics **11** 87-93.

Coupland G (1992) Transposon tagging in *Arabidopsis*. *In* Methods in *Arabidopsis* research. (Editors, Koncz C, Chua N-H, Schell J) World Scientific, New Jersey. 290-309.

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

Deroles S, Gardner R (1988a) Expression and inheritance of kanamycin resistance in a large number of transgenic petunias generated by *Agrobacterium*-mediated transformation. Plant Molecular Biology **11** 355-364.

Deroles S, Gardner R (1988b) Analysis of the T-DNA structure in a large number of transgenic petunias generated by *Agrobacterium*-mediated transformation. Plant Molecular Biology **11** 365-377.

Ditta G, Stanfield S, Corbin D, Helinski DR (1980) Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti.* Proceedings of the National Academy of Sciences, USA **77** 7347-7351.

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin **19** 11-15.

Dretzen G, Bellard M, Sassone-Corsi P, Chambon P (1981) A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. Analytical Biochemistry **112** 295-298.

Errington J (1990) A rapid and reliable one-step method for isolating DNA fragments from agarose gels. Nucleic Acids Research **18** 5324.

Feldman K, Marks D, Christianson M, Quatrano R (1989) A dwarf mutant of *Arabidopsis* generated by T-DNA insertion mutagenesis. Science **243** 1351-1354.

Feldmann K (1991) T-DNA insertion mutagenesis in *Arabidopsis:* Mutational spectrum. The Plant Journal **1** 71-82.

Feldmann K (1992) T-DNA insertion mutagenesis in *Arabidopsis:* Seed infection-transformation. *In* Methods in *Arabidopsis* research. (Editors, Koncz C, Chua N, Schell J) World Scientific, New Jersey. 274-289.

Feldmann K, Carlson T, Coomber S, Farrance C, Mandel M, Wierzbicki M (1990) T-DNA insertional mutagenesis in *Arabidopsis thaliana. In* Horticultural biotechnology (Editors, Bennett AB, O'Neil SD) Wiley-Liss, New York. 109-120.

Feldmann K, Marks M (1987) *Agrobacterium*-mediated transformation of germinating seeds of *Arabidopsis thaliana:* A non-tissue culture approach. Molecular and General Genetics **208** 1-9.

Figurski DH, Helinski DR (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. Proceedings of the National Academy of Sciences, USA **76** 1648-1652.

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

Fobert PR, Miki B, Iyer V (1991) Detection of gene regulatory signals in plants revealed by T-DNA-mediated fusions. Plant Molecular Biology **17** 837-851.

Fobert PR, Labbe H, Cosmopoulos J, Gottlob-McHugh S, Ouellet T, Hattori J, Sunohara G, Iyer V, Miki B (1994) T-DNA tagging of a seed coat-specific cryptic promoter in tobacco. The Plant Journal **6** 567-577.

Fors L, Saavedra RA, Hood L (1990) Cloning of the shark Po promoter using a genomic walking technique based on the polymerase chain reaction. Nucleic Acids Research **18** 2793-2799.

Forsthoefel N, Wu Y, Schulz B, Bennett M, Feldmann K (1992) T-DNA insertion mutagenesis in *Arabidopsis:* Prospects and perspectives. Australian Journal of Plant Physiology **19** 353-366.

Frisch, DA, Harris-Haller LW, Yokubaitis NT, Thomas TL, Hardin SG, Hall TC (1995) Complete sequence of the binary vector Bin 19. Plant Molecular Biology **27** 405-409.

Fritze K, Czaja I, Walden R (1995) T-DNA tagging of genes influencing polyamine

metabolism: Isolation of mutant plant lines and rescue of DNA promoting growth in the presence of a polyamine biosynthetic inhibitor. The Plant Journal **7** 261-271.

Gamborg OL, Miller RA, Ojima K (1968)Nutrient requirements of suspension cultures of soybean root cells. Experimental Cell Research **50** 151-158.

Gardiner K, Mural R (1995) Getting the message: Identifying transcribed sequences. Trends in Genetics **11** 77-79.

Gheysen G, Villarroel R, Van Montagu M (1991) Illegitimate recombination in plants: A model for T-DNA integration. Genes and Development **5** 287-297.

Gierl A, Saedler H (1992) Plant-transposable elements and gene tagging. Plant Molecular Biology **19** 39-49.

Goddijn OJM, Lindsey K, van der Lee FM, Klap JC, Sijmons PC (1993) Differential gene expression in nematode-induced feeding structures of transgenic plants harbouring promoter-*gus*A fusion constructs. The Plant Journal **4** 863-873.

Goldsbrough A, Bevan M (1991) New patterns of gene activity in plants detected using an *Agrobacterium* vector. Plant Molecular Biology **16** 263-269.

Gossler A, Zachgo J (1993) Gene and enhancer trap screens in ES cell chimeras. *In* Gene targeting: A practical approach. (Editor, Joyner A) IRL Press, New York. 181-227.

Grant SGN, Jessee J, Bloom FR, Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proceedings of the National Academy of Sciences, USA **87** 4645-4649.

Grevelding C, Fantes V, Kember E, Schell J, Masterson R (1993) Single-copy T-DNA insertions in *Arabidopsis* are the predominant form of integration in root-derived transgenics, whereas multiple insertions are found in leaf discs. Plant Molecular Biology **23** 847-860.

Gridley T, Soriano P, Jaenisch R (1987) Insertional mutagenesis in mice. Trends in Genetics **3** 162-166.

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. Journal of Molecular Biology **166** 557-580.

Hayashi H, Czaja I, Lubenow H, Schell J, Walden R (1992) Activation of a plant gene by T-DNA tagging: Auxin-independent growth in vitro. Science **258** 1350-1355.

Hehl R (1994) Transposon tagging in heterologous host plants. Trends in Genetics **10** 385-386.

Hepburn AG, Belanger FC, Mattheis JR (1987) DNA methylation in plants. Developmental Genetics **8** 475-493.

Herman L, Van Montagu M, Depicker A (1986) Isolation of tobacco DNA segments with plant promoter activity. Molecular and Cellular Biology **6** 4486-4492.

Herman PL, Marks MD (1989) Trichome development in *Arabidopsis thaliana*. II. Isolation and complementation of the GLABROUS1 gene. The Plant Cell **1** 1051-1055.

Hobbs SLA, Warkentin TD, DeLong CMO (1993) Transgene copy number can be positively or negatively associated with transgene expression. Plant Molecular Biology **21** 17-26.

Hobbs SLA, Kpodar P, DeLong CMO (1990) The effect of T-DNA copy number, position and methylation on reporter gene expression in tobacco transformants. Plant Molecular Biology **15** 851-864.

Hoekema A, Hirsch PR, Hooykaas PJJ, Schilperoort RA (1983). A binary plant vector strategy based on separation of *vir*- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. Nature **303** 179-180.

Holmes DS, Quigley M (1981) A rapid boiling method for the preparation of bacterial plasmids. Analytical Biochemistry **114** 193-197.

Hooykaas P, Schilperoort R (1992) *Agrobacterium* and plant genetic engineering. Plant Molecular Biology **19** 15-38.

Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT (1985) A simple and general method for transferring genes into plants. Science **227** 1229-1231.

Ingelbrecht I, Breyne P, Vancompernolle K, Jacobs A, Van Montagu M, Depicker A (1991) Transcriptional interference in transgenic plants. Gene **109** 239-242.

Janssen B, Gardner R (1989) Localized transient expression of GUS in leaf discs following cocultivation with *Agrobacterium*. Plant Molecular Biology **14** 61-72.

Jefferson RA (1987) Assaying chimeric genes in plants: The GUS gene fusion system. Plant Molecular Biology Reporter **5** 387-405.

Jiang C, Langridge W, Szalay A (1992) Identification of plant genes *in vivo* by tagging with T-DNA border-linked luciferase genes followed by inverse polymerase chain reaction amplification. Plant Molecular Biology Reporter **10** 345-361.

Jones PA (1985) Altering gene expression with 5-azacytidine. Cell **40** 485-486.

Jorgensen R (1990) Altered gene expression in plants due to *trans* interactions between homologous genes. Trends in Biotechnology **8** 340-344.

Kado C (1991) Molecular mechanisms of crown gall tumorigenesis. Critical Reviews in Plant Science **10** 1-32.

Keller B, Lamb CJ (1989) Specific expression of a novel cell wall hydroxyproline-rich glycoprotein gene in lateral root initiation. Genes and Development **3** 1639-1646.

Kerr D, Griffiths A, White D (1991) Gene tagging to identify novel plant promoters. Queenstown Molecular Meeting Presentation. August 19-23, Queenstown, New Zealand.

Kertbundit S, De Greve H, Deboeck F, Van Montagu M, Hernalsteens J (1991) In *vivo* random β-glucuronidase gene fusions in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences, USA **88** 5212-5216.

Kleckner N (1981) Transposable elements in prokaryotes. Annual Review of Genetics **15** 341-404.

Koltunow AM, Truettner J, Cox KH, Wallroth M, Goldberg RG (1990) Different temporal and spatial gene expression patterns occur during anther development. The Plant Cell **2** 1201-1224.

Koncz C, Nemeth K, Redei G, Schell J (1992) T-DNA insertional mutagenesis in *Arabidopsis*. Plant Molecular Biology **20** 963-976.

Koncz C (1989) Unpublished notes. Max Planck Institute, Cologne. 75p.

Koncz C, Martini N, Mayerhofer R, Koncz-Kanman Z, Korber H, Redei G, Schell J (1989) High-frequency T-DNA-mediated gene tagging in plants. Proceedings of the National Academy of Sciences, USA **86** 8467-8471.

Koncz C, Schell J (1986) The promoter of the T_L-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. Molecular and General Genetics **204** 383-396.

Kuhlemeier C, Green P, Chua N (1987) Regulation of gene expression in higher plants. Annual Review of Plant Physiology **38** 221-257.

Kuhlemeier C (1992) Transcriptional and post-transcriptional regulation of gene expression in plants. Plant Molecular Biology **19** 1-14.

Lindsey K, Wei W, Clarke M, McArdle H, Rooke L, Topping J (1993) Tagging genomic sequences that direct transgene expression by activation of a promoter trap in plants. Transgenic Research **2** 33-47.

Loh EY, Elliot JF, Cwirla S, Lanir LL, Davis MM (1989) Polymerase chain reaction with single sided specificity: Analysis of T cell receptor δ chain. Science **243** 217-220.

Lorow D, Jessee J (1990) MAX efficiency DH10B: a new host for cloning methylated DNA. Focus **12** 19-20.

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory. 545p.

Martin T, Wohner R-V, Hummel S, Willmitzer L, Frommer WB (1992) The GUS reporter system as a tool to study plant gene expression. *In* GUS protocols: Using the GUS gene as a reporter of gene expression. (Editor Gallagher SR) Academic Press, San Diego. 23-43.

Martineau B, Voelker TA, Sanders RA (1994) On defining T-DNA. The Plant Cell **6** 1032-1033.

Marton L, Hrouda M, Pecsvaradi A, Czako M (1994) T-DNA-insert-independent mutations induced in transformed plant cells during *Agrobacterium* co-cultivation. Transgenic Research **3** 317-325.

Matassi G, Melis R, Macaya G, Bernardi G (1991) Compositional bimodality of the nuclear genome of tobacco. Nucleic Acids Research **19** 5561-5567.

Matsumoto S, Ito Y, Hosoi T, Takahashi Y, Machida Y (1990) Integration of *Agrobacterium* T-DNA into a tobacco chromosome: Possible involvement of DNA homology between T-DNA and plant DNA. Molecular and General Genetics **224** 309-316.

Matzke MA, Matzke AJM (1991) Differential inactivation and methylation of a transgene in plants by two suppressor loci containing homologous sequences. Plant Molecular Biology **16** 821-830.

Mayerhofer F, Koncz-Kalman Z, Nawrath C, Bakkeren G, Crameri A, Angelis K, Redei G, Schell J, Hohn B, Koncz C (1991) T-DNA integration: A mode of illegitimate recombination in plants. EMBO Journal **10** 697-704.

Meyer P, Niedenhof I, ten Lohuis M (1994) Evidence for cytosine methylation of non-symmetrical sequences in transgenic *Petunia hybrida.* EMBO Journal **13** 2084-2088.

Meyer P, Heidmann I (1994) Epigenetic variants of a transgenic petunia line show hypermethylation in transgene DNA: An indication for specific recognition of foreign DNA in transgenic plants. Molecular and General Genetics **243** 390-399.

Meyer P, Linn F, Heidmann I, Meyer zA H, Niedenhof I, Saedler H (1992) Endogenous and environmental factors influence 35S promoter methylation of a maize A1 gene construct in transgenic petunia and its colour phenotype. Molecular and General Genetics **231** 345-352.

Meyerowitz E (1992) Introduction to the *Arabidopsis* genome. *In* Methods in *Arabidopsis* research. (Editors, Koncz C, Chua N-H, Schell J) World Scientific, New Jersey. 100-118.

Mueller P, Wold B (1989) In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. Science **246** 780-786.

Mueller P, Wold B (1990) Ligation mediated single-sided PCR for genomic sequencing and footprinting. Unpublished notes. California Institute of Technology. 13p.

Muller E, Brown PTH, Hartke S, Lorz H (1990) DNA variation in tissue-culture-derived rice plants. Theoretical and Applied Genetics **80** 673-679.

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiologia Plantarum **15** 473-497.

Ochman H, Gerber AS, Hartl DL (1988) Genetic applications of an inverse polymerase chain reaction. Genetics **120** 621-623.

Okamuro JK, Goldberg RB (1985) Tobacco single-copy DNA is highly homologous to sequences present in the genomes of its diploid progenitors. Molecular and General Genetics **198** 290-298.

Ooms G, Hooykaas PJJ, Van Veen RJM, Van Beelen P, Regensburg-Tuink TJG, Schilperoort RA (1982). Octopine Ti-plasmid deletion mutants of *Agrobacterium* tumefaciens with emphasis on the right side of the T-region. Plasmid **7** 15-29.

Ott RW, Chua N-H (1990) Enhancer sequences from *Arabidopsis thaliana* obtained by library transformation of *Nicotiana tabacum*. Molecular and General Genetics **223** 169-179.

Paszkowski J, Saul MW (1986) Direct gene transfer to plants. Methods In Enzymology **118** 668-684.

Peralta E, Hellmiss R, Ream W (1986) *Overdrive*, a T-DNA transmission enhancer on the *A. tumefaciens* tumour-inducing plasmid. EMBO Journal **5** 1137-1142.

Ponce MR, Micol JL (1992) PCR amplification of long DNA fragments. Nucleic Acids Research **20** 623.

Promega Corporation (1991) Promega protocols and applications guide. Second edition. Promega Corporation, Madison. 422p.

Rago R, Mitchen J, Wilding G (1990) DNA fluorometric assay in 96-well tissue culture plates using hoechst 33258 after cell lysis by freezing in distilled water. Analytical Biochemistry **191** 31-34.

Ream W (1989) *Agrobacterium tumefaciens* and interkingdom genetic exchange. Annual Review of Phytopathology **27** 583-618.

Robinson N, Shirsat A, Gatehouse J (1993) Regulation of gene expression. *in* Plant biochemistry and molecular biology. (Editors, Lea P, Leegood R) John Wiley & Sons Ltd, England. 222-240.

Saiki R, Gelfand D, Stoffel S, Scharf S, Higuchi R, Horn G, Mullis K, Erlich H (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239** 487-491.

Sawadogo M, Van Dyke MW (1991) A rapid method for the purification of deprotected oligodeoxynucleotides. Nucleic Acids Research **19** 674.

Selker EU (1990) DNA methylation and chromatin structure: A view from below. Trends in Biochemical Sciences **15** 103-107.

Shields R (1993) Transposons hop out of maize. Nature **363** 669-670.

Simon R, Priefer U, Puhler A (1983) A broad host range mobilization system for *in vivo* genetic engineering: Transposon mutagenesis in gram negative bacteria. Bio/Technology **1** 784-791.

Smith DL, Fedoroff NV (1995) *LRP1*, a gene expressed in lateral and adventitious root primordia of Arabidopsis. The Plant Cell **7** 735-745.

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of Molecular Biology **98** 503-517.

Springer P, McCombie R, Sundaresan V, Martienssen R (1995) Gene trap tagging of *prolifera*, an essential MCM2-3-5-like gene in *Arabidopsis*. Science **268** 877-880.

Stachel SE, An G, Flores C, Nester EW (1985) A Tn3 *lacZ* transposon for the random generation of β-galactosidase gene fusions: Application to the analysis of gene expression in *Agrobacterium*. EMBO Journal **4** 891-898.

Suntio T, Teeri T (1994) A new bifunctional reporter gene for *in-vivo* tagging of plant promoters. Plant Molecular Biology Reporter **12** 43-57.

Taylor BH, Scheuring CF (1994) A moleular marker for lateral root initiation: The *RSI-1* gene of tomato (*Lycopersicon esculentum* Mill) is activated in early lateral root primordia. Molecular and General Genetics **243** 148-157.

Teeri TH, Herrera-Estrella L, Depicker A, Van Montagu M, Palva ET (1986) Identification of plant promoters *in situ* by T-DNA-mediated transcriptional fusions to the *npt*II gene. EMBO Journal **5** 1755-1760.

Teeri TH, Lehvaslaiho H, Franck M, Uotila J, Heino P, Palva ET, Van Montagu M, Herrera-Estrella L (1989) Gene fusion to *lacZ* reveal new expression patterns of chimeric genes in transgenic plants. EMBO Journal **8** 343-350.

Thomas CM, Jones DA, English JJ, Carroll BJ, Bennetzen JL, Harrison K, Burbidge A, Bishop GJ, Jones JDG (1994) Analysis of the chromosomal distribution of

transposon-carrying T-DNAs in tomato using the inverse polymerase chain reaction. Molecular and General Genetics **242** 573-585.

Topping F, Agyeman F, Henricot B, Lindsey K (1994) Identification of molecular markers of embryogenesis in *Arabidopsis thaliana* by promoter trapping. Plant Journal **5** 895-903.

Triglia T, Peterson M, Kemp D (1988) A procedure for *in vitro* amplification of DNA segments that lie outside the boundaries of known sequence. Nucleic Acids Research **16** 8186.

United States Biochemical (1992) 'Sequenase' version 2.0 DNA sequencing kit. United States Biochemical, Cleveland. 25 p.

Valvekens D, Van Montagu M, Van Lijsebettens M (1988) *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. Proceedings of the National Academy of Sciences, USA **85** 5536-5540.

van den Berg C, Willemsen V, Hage W, Weisbeek P, Scheres B (1995) Cell fate in the *Arabidopsis* root meristem determined by directional signalling. Nature **378** 62-65.

Van Lijsebettens M, Vanderhaeghen R, Van Montagu M (1991) Insertional mutagenesis in *Arabidopsis thaliana:* Isolation of a T-DNA-linked mutation that alters leaf morphology. Theoretical and Applied Genetics **81** 277-284.

Van Larabeke N, Engler G, Holsters M, Van den Elsacker S, Zaenen J, Schilperoort RA, Schell J (1974) Large plasmid in *Agrobacterium tumefaciens* essential for crown gall-inducing ability. Nature **252** 169-170.

Vieira J, Messing J (1982) The pUC plasmids, and M13mp&-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19** 259-268.

Virts E, Gelvin S (1985) Analysis of transfer of tumor-inducing plasmids from *Agrobacterium tumefaciens* to *Petunia* protoplasts. Journal of Bacteriology **162** 1030-1038.

Walbot V (1992) Strategies for mutagenesis and gene cloning using transposon tagging and T-DNA insertional mutagenesis. Annual Review of Plant Physiology **43** 49-82.

Walden R, Schell J (1990) Techniques in plant molecular biology - progress and problems. European Journal of Biochemistry **192** 563-576.

Walden R, Fritze K, Hayashi H, Miklashevichs E, Harling H, Shell J (1994a) Activation tagging: A means of isolating genes implicated as playing a role in plant growth and development. Plant Molecular Biology **26** 1521-1528.

Walden R, Hayashi H, Lubenow H, Czaja I, Schell J (1994b) Auxin inducibility and developmental expression of *axi* 1: A gene directing auxin independent growth in tobacco protoplasts. EMBO Journal **13** 4729-4736.

Wang K, Herrera-Estrella L, Van Montagu M, Zambryski P (1984) Right 25 bp terminus sequence of the nopaline T-DNA is essential for and determines direction of DNA transfer from *Agrobacterium* to the plant genome. Cell **38** 455-462.

Waugh R, Brown J (1990) Plant gene structure and expression. *in* Plant genetic engineering. (Editor, Grierson D) Chapman and Hall, New York. 1-37.

Weber H, Ziechmann C, Graessmann A (1990) *In vitro* DNA methylation inhibits gene expression in transgenic tobacco. EMBO Journal **9** 4409-4415.

Westfall B, Sitaraman K, Berninger M, Mertz LM (1995) Elongase reagents for amplification of long DNA templates. Focus **17** 62-65.

Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33** 103-119.

Yanofsky MF, Ma H, Bowman JL, Drews GN, Feldmann KA, Meyerowitz EM (1990) The protein encoded by the *Arabidopsis* homeotic gene agamous resembles transcription factors. Nature **346** 35-39.

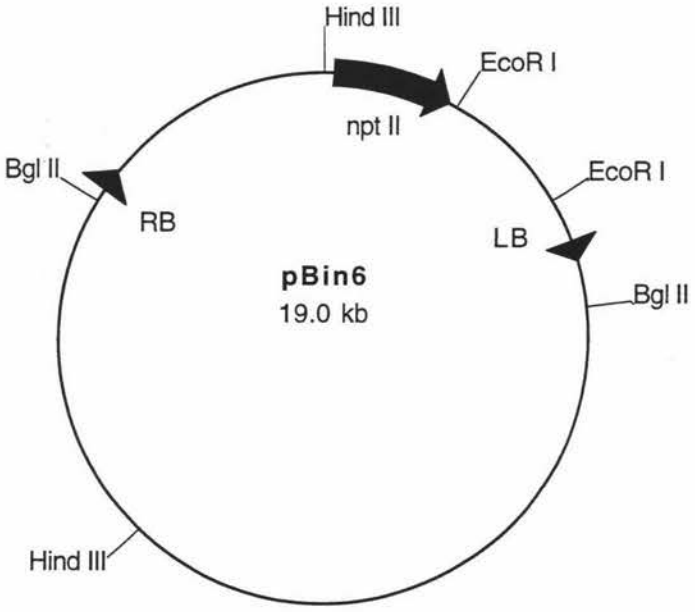
Zambryski P, Holsters M, Kruger K, Depicker A, Schell J, Van Montagu M, Goodman HM (1980) Tumor DNA structure in plant cells transformed by *A. tumefaciens*. Science **209** 1385-1391.

Zhixing Y, Nahon J-L (1995) DNA gyrase improves DNA transformation of *E. coli* cells with large recombinant plasmids. Nucleic Acids Research **23** 3353-3354.

8.0 Appendix

8.1 Plasmids

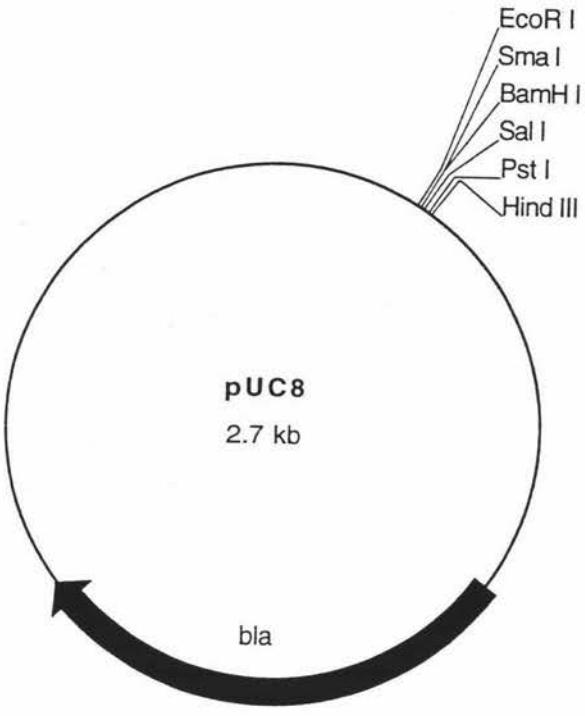
8.1.1 Physical map of pBin6



Information from Bevan (1984) and this study.

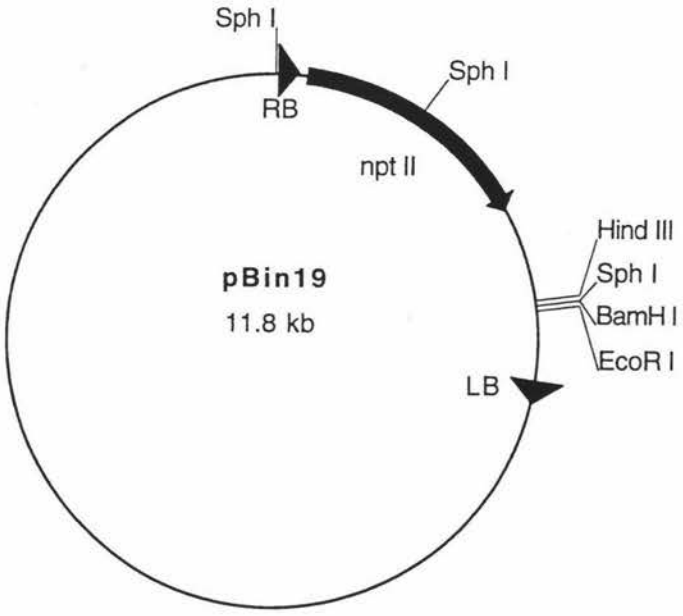
Abbreviations: LB, left border of T-DNA; RB, right border of T-DNA; *npt*II, chimaeric Pnos−*npt*II−3′*nos* gene.

8.1.2 Physical map of pUC8



Cloning plasmid with polylinker within β-galactosidase α-peptide coding sequence. Abbreviation: *bla*, β-lactamase gene.

8.1.3 Physical map of pBin19



Information from Bevan (1984) and Frisch *et al.* (1995).

Abbreviations: RB, right border; LB, left border; *npt*II, chimaeric Pnos−*npt*II−3′*nos* gene.

8.2 Southern hybridisations

Summary of autoradiographs

Plant	gel	*Hin*dIII lane	*Eco*RI lane	number of significant bands hybridising with probe/enzyme combinations				estimate of T-DNA fragments[1]	
				gus/ *Hin*dIII	*gus*/ *Eco*RI	*npt*II/ *Hin*dIII	*npt*II/ *Eco*RI	left border	right border
W23.1	2	1	11	2	4	3	1	3	4
W23.2	2	2	12	2	1	3	2	3	2
W47.2	2	3	13	2	6	5	1		
W47.3	2	4	14	1	1	1	1	1	1
W47.4	2	5	15	3	3	3	2	3	3
A26.6	2	6	16	4	5	6	2	6	5
A18.2	2	7	17	3	5	4	2	4	5
A41.1	2	8	18	1	2	2	1	2	2
A53.1	2	9	19	4	3	5	1	5	4
A82.1	2	10	20	4	5	3	1	3	5
W1.1	4	1	11	nd[2]	nd	nd	2		
W5.1	4	2	12	4	4	5	3	3	4
W5.2	4	3	13	3	3	3	1	3	3
W5.3	4	4	14	2	nd	1	nd		
W5.4	4	5	15	2	3	2	2	2	3
W5.5	4	6	16	7	5	6	5	4	5
W5.6	4	7	17	2	4	1	2	1	4
W5.7	4	8	18	3	nd	3	nd		
W5.8	4	9	19	2	3	3	2	3	3
W5.9	4	10	20	1	1	2	2	2	1
W2.2	5	1	11	1	1	2	2		
W2.3	5	2	12	3	3	4	3	3	3
W2.4	5	3	13	1	1	1	0	1	1
W2.5	5	4	14	nd	3	nd	2		
W2.6	5	5	15	3	3	2	2	2	3
W3.1	5	6	16	1	1	1	1	1	1
W4.1	5	7	17	3	2	3	2	2	3
W4.2	5	8	18	nd	1	3	2		
W10.1	5	9	19	nd	1	nd	nd		
W10.2	5	10	20	1	1	1	nd		

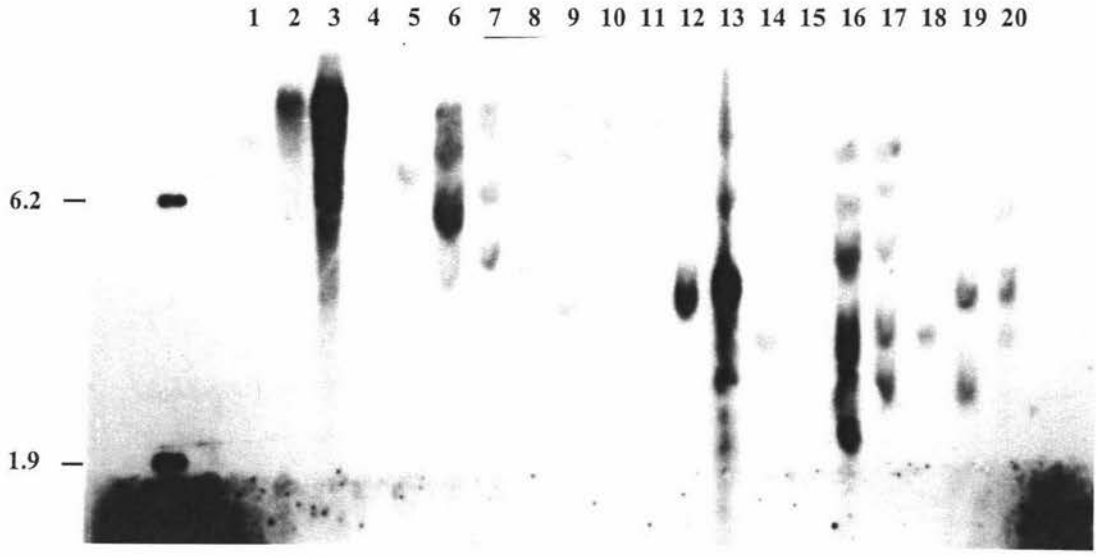
Plant	gel	*Hin*dIII lane	*Eco*RI lane	number of significant bands hybridising with probe/enzyme combinations				estimate of T-DNA fragments	
				gus/ *Hin*dIII	*gus*/ *Eco*RI	*npt*II/ *Hin*dIII	*npt*II/ *Eco*RI	left border	right border
W6.1	6	1	11	4	3	3	3	2	4
W6.2	6	2	12	2	1	1	1	1	2
W6.3	6	3	13	8	6	6	6	6	8
W6.4	6	4	14	0	0	1	1	1	0
W8.1	6	5	15	3	2	2	3	1	2
W8.2	6	6	16	1	1	1	1	1	1
W8.4	6	7	17	1	3	5	4	4	3
W9.1	6	8	18	3	2	2	2	2	2
W11.1	6	9	19	3	3	2	2	2	3
W12.1	6	10	20	3	3	3	4	2	2
W14.1	7	1	11	1	1	2	3	2	1
W16.1	7	2	12	4	6	3	2	3	6
W18.1	7	3	13	4	4	4	5	4	4
W19.1	7	4	14	2	3	4	5	4	3
W20.1	7	5	15	3	1	2	1	2	3
W22.1	7	6	16	3	2	4	2	4	3
W24.1	7	7	17	0	0	1	1	1	0
W25.1	7	8	18	1	1	1	1	1	1
W26.1	7	9	19	7	6	4	4	4	6
W26.2	7	10	20	3	2	3	7	3	3
W27.1	8	1	11			nd	2		
W28.1	8	2	12			5	2		
W29.1	8	3	13			4	2		
W29.2	8	4	14			0	0		
W30.1	8	5	15			1	1		
W31.1	8	6	16			2	1		
W34.1	8	7	17			2	1		
W39.1	8	8	18			3	1		
W40.1	8	9	19			1	1		
W40.2	8	10	20			1	nd		

Plant	gel	*Hin*dIII lane	*Eco*RI lane	number of significant bands hybridising with probe/enzyme combinations				estimate of T-DNA fragments	
				gus/ *Hin*dIII	*gus*/ *Eco*RI	*npt*II/ *Hin*dIII	*npt*II/ *Eco*RI	left border	right border
W41.1	9	1	11	2	2	3	1		
W42.2	9	2	12	2	2	2	1		
W43.1	9	3	13	2	2	3	2		
W43.2	9	4	14	4	6	6	2		
W43.4	9	5	15	5	5	8	3		
W43.6	9	6	16	nd	nd	1	1		
W48.1	9	7	17	4	4	4	2		
W48.2	9	8	18	nd	nd	1	1		
W48.5	9	9	19	nd	nd	nd	nd		
W48.6	9	10	20	3	2	4	1		
A26.6	10	1	2			5	4		
A88.1	10	4	5			4	3		
W23.2	10	7	8			4	5		
W50.1	10	10	11			4	4		
W50.2	10	13	14			7	5		
W37.1	11	1	11	nd	1				
W37.2	11	2	12	nd	2				
W7.3	11	3	13	nd	nd				
W46.1	11	4	14	2	2				
W46.2	11	5	15	nd	nd				
W46.3	11	6	16	1	nd				
W46.4	11	7	17	1	nd				
W45.1	11	8	18	nd	nd				
W45.2	11	9	19	2	2				
W45.3	11	10	20	nd	nd				

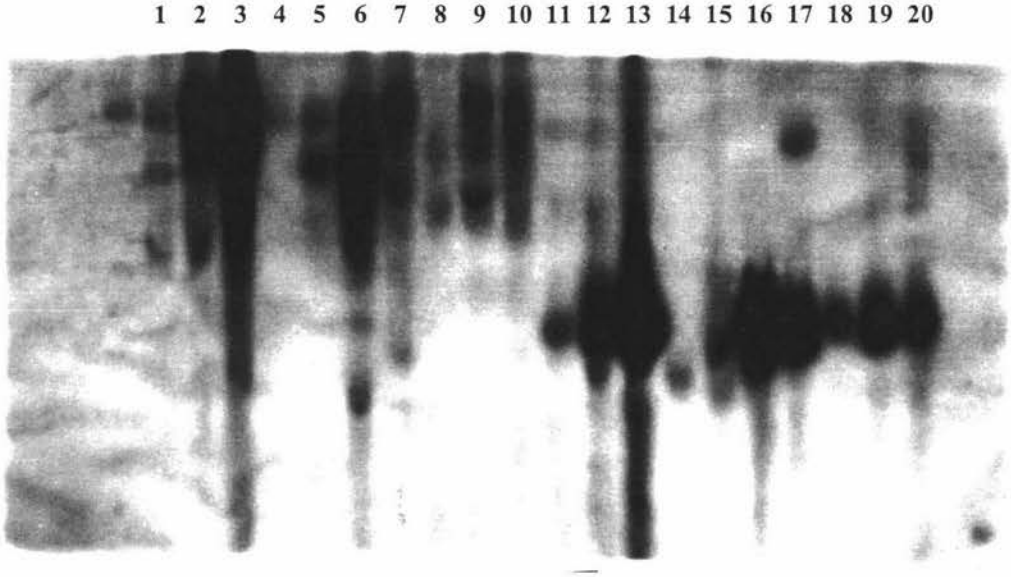
Plant	gel	*Hin*dIII lane	*Eco*RI lane	number of significant bands hybridising with probe/enzyme combinations				estimate of T-DNA fragments	
				gus/ *Hin*dIII	*gus*/ *Eco*RI	*npt*II/ *Hin*dIII	*npt*II/ *Eco*RI	left border	right border
W51.3	12	1	11	nd	2				
W52.1	12	2	12	nd	2				
W52.2	12	3	13	nd	nd				
W52.3	12	4	14	nd	nd				
W53.1	12	5	15	nd	nd				
W53.2	12	6	16	nd	1				
W44.1	12	7	17	5	7				
Control	12	8	18	0	0				
W43.4	12	9	19	7	1				
W27.1	12	10	-	1	-				
W40.2	12	-	20	-	nd				

[1] determined after eliminating potential partial bands.
[2] not determined.

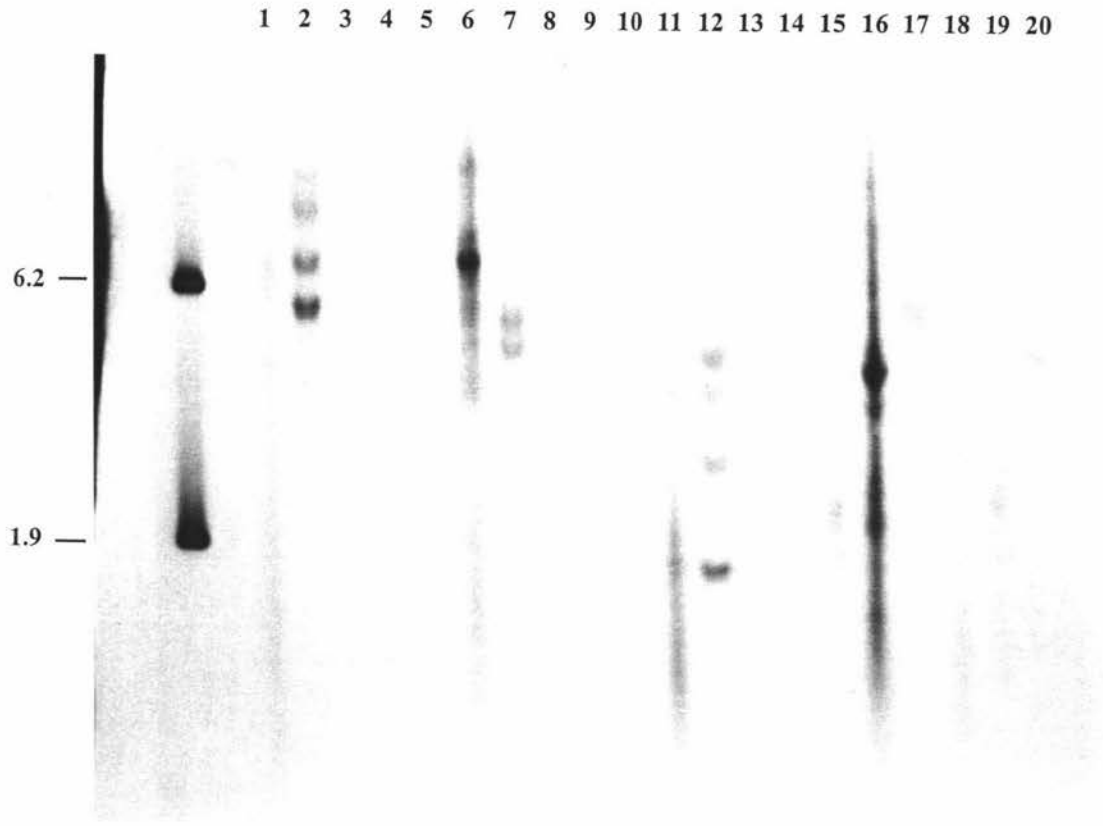
gel 2: *gus* probed



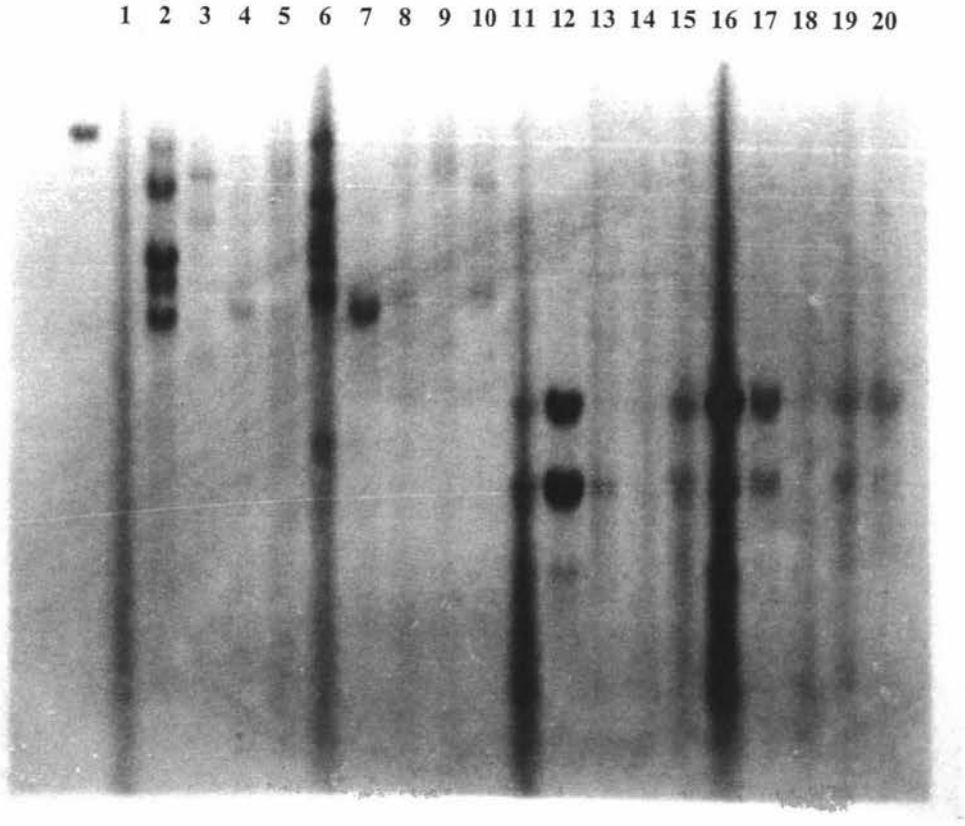
gel 2: *npt* II probed



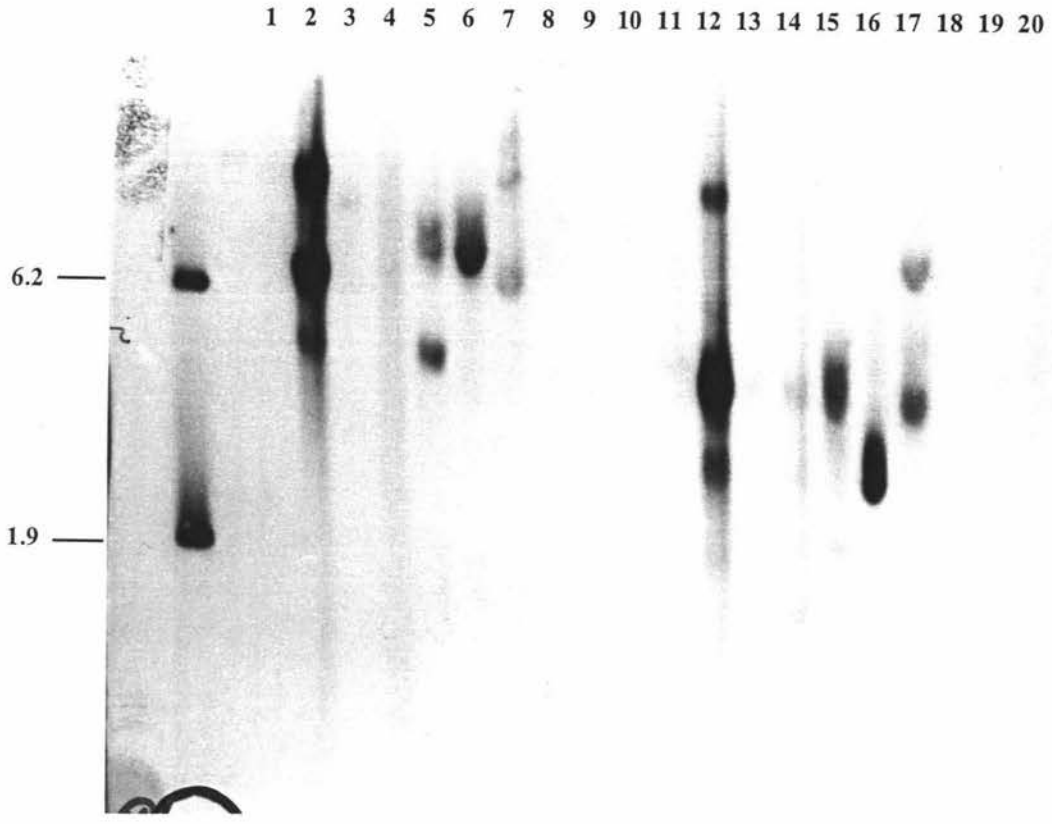
gel 4: *gus* probed



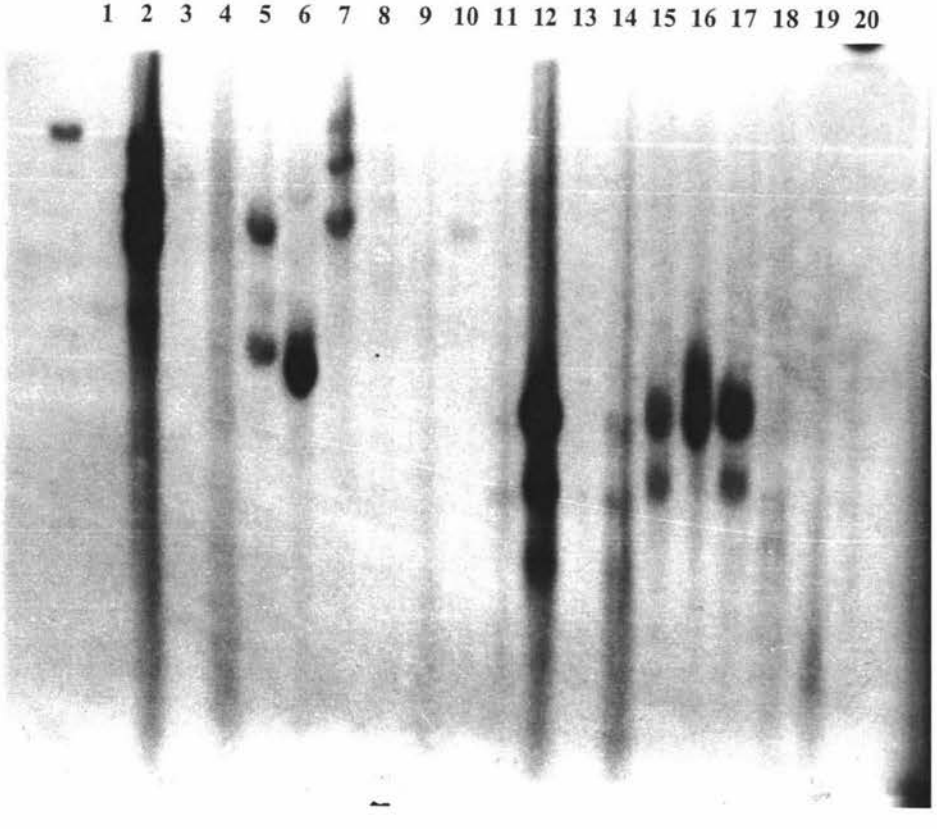
gel 4: *npt* II probed



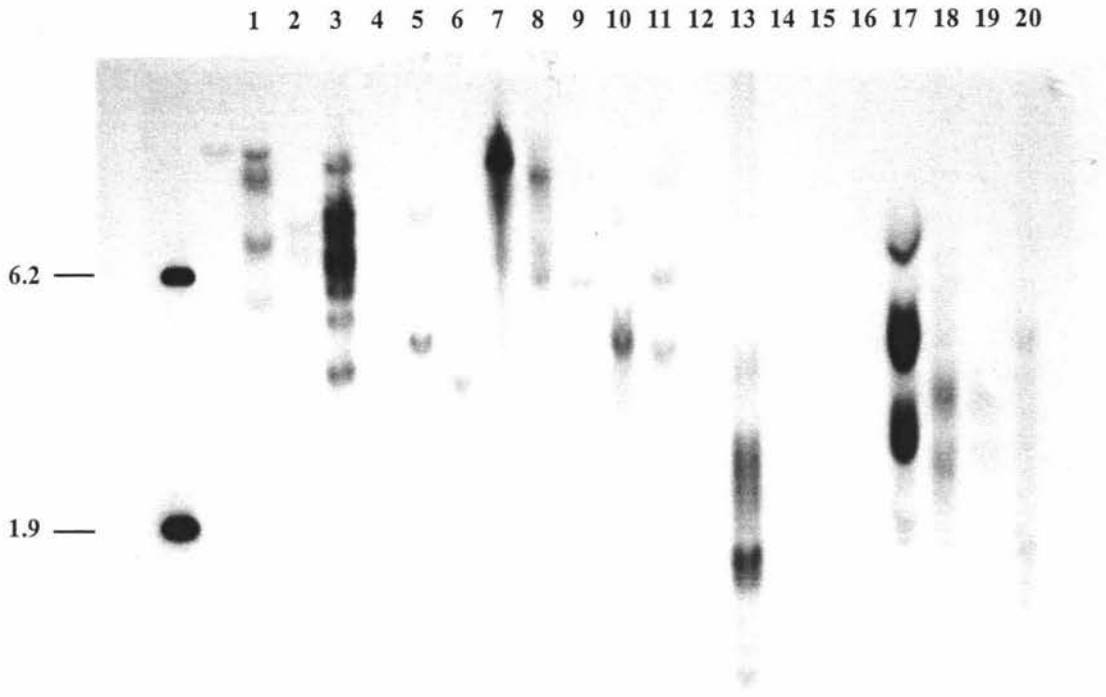
gel 5: *gus* probed



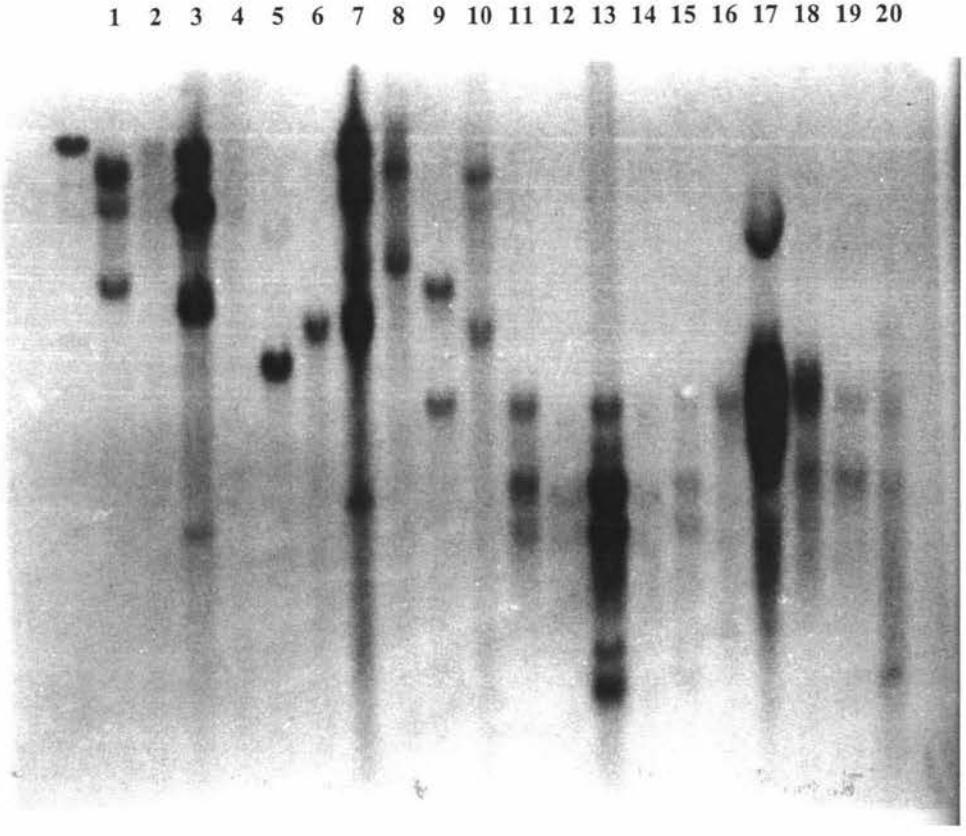
gel 5: *npt* II probed



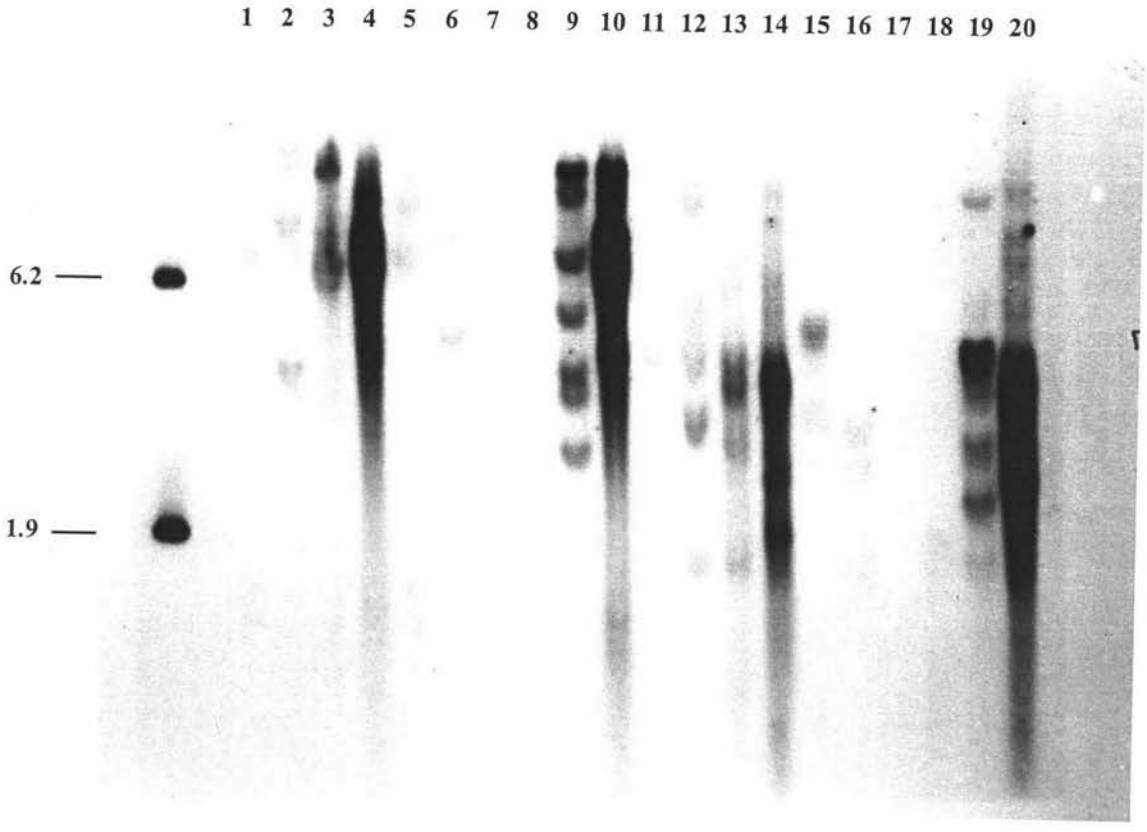
gel 6: *gus* probed



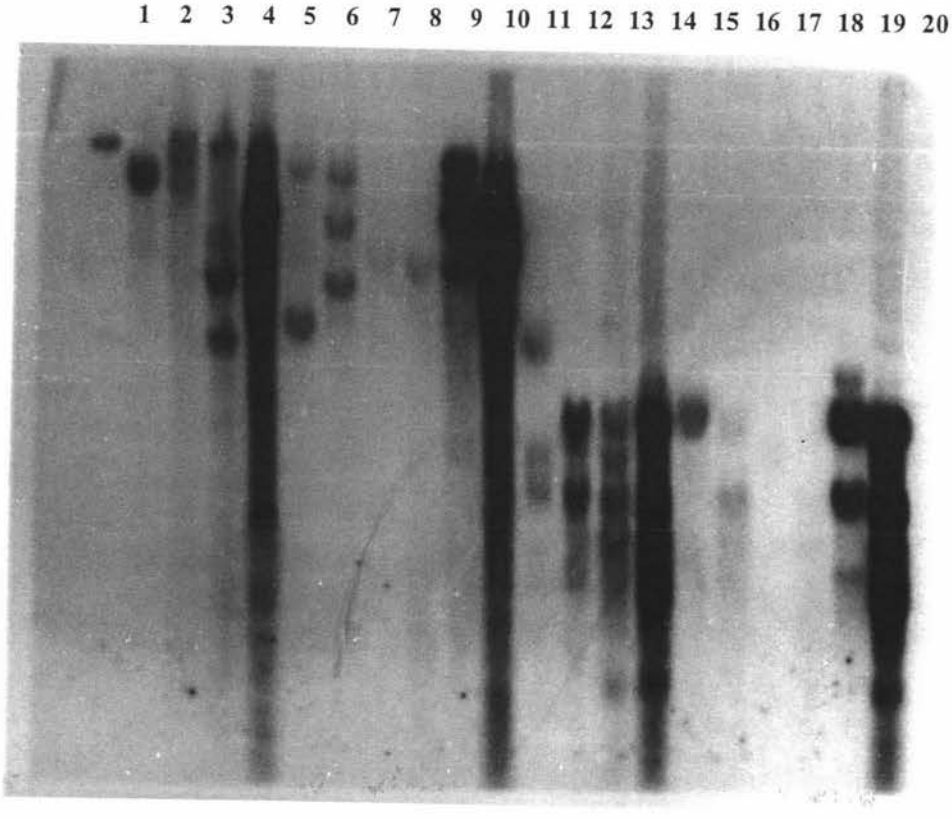
gel 6: *npt* II probed



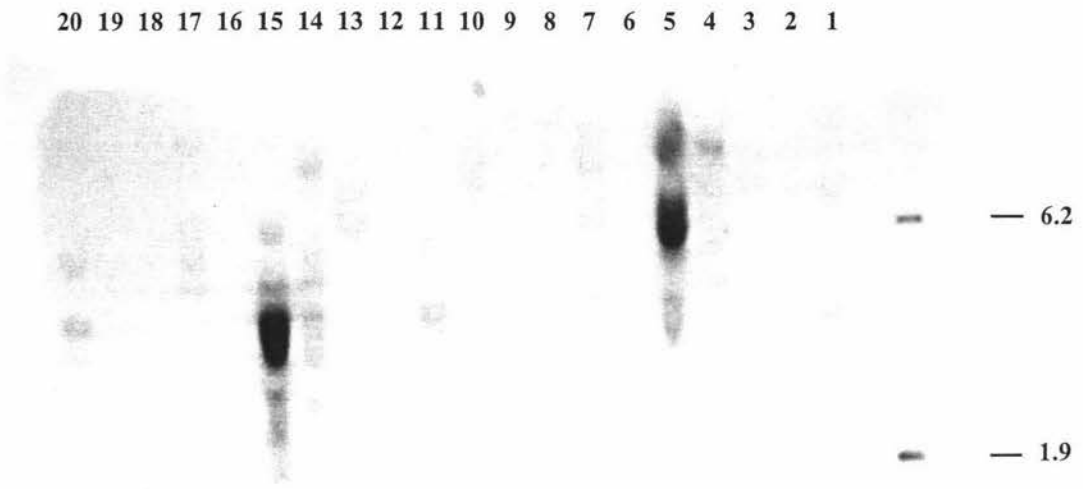
gel 7: *gus* probed



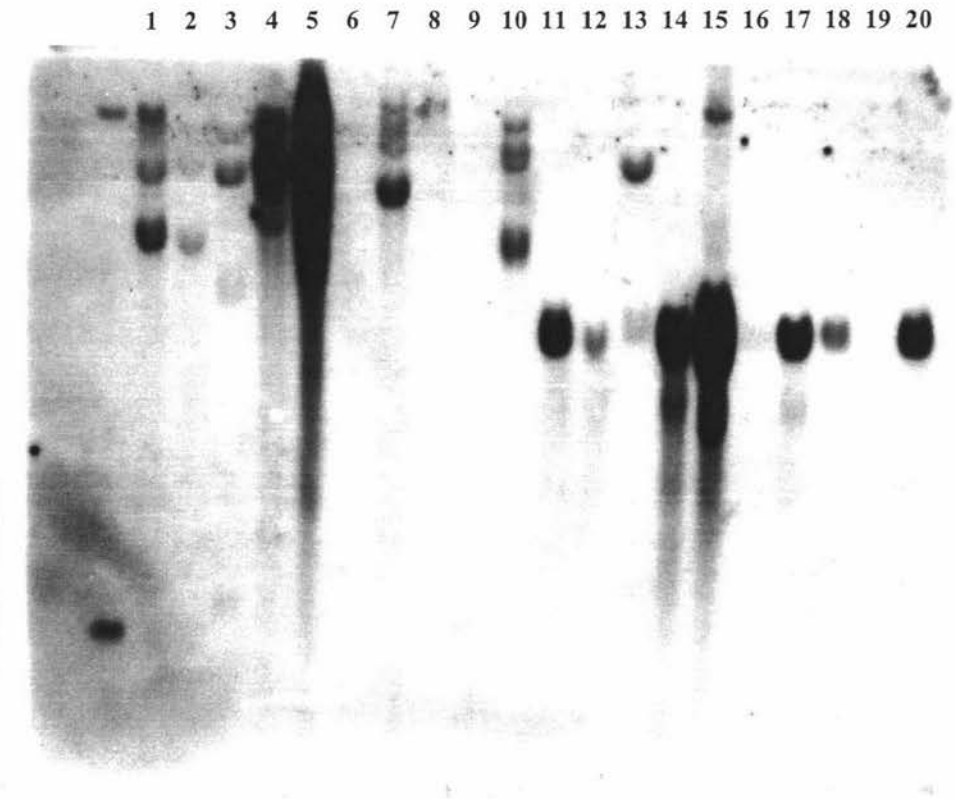
gel 7: *npt* II probed



gel 9: *gus* probed

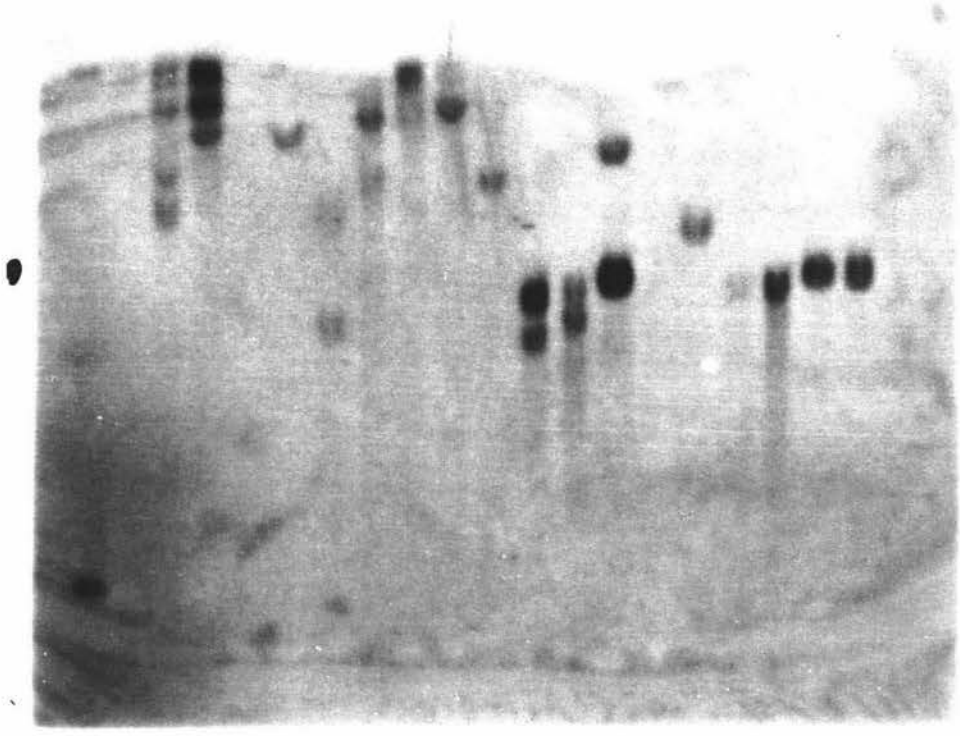


gel 9: *npt* II probed



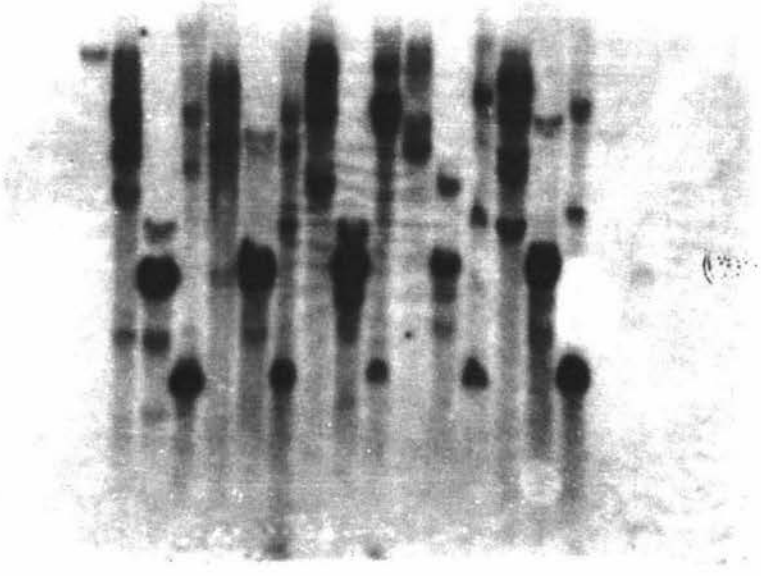
gel 8: *npt* II probed

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

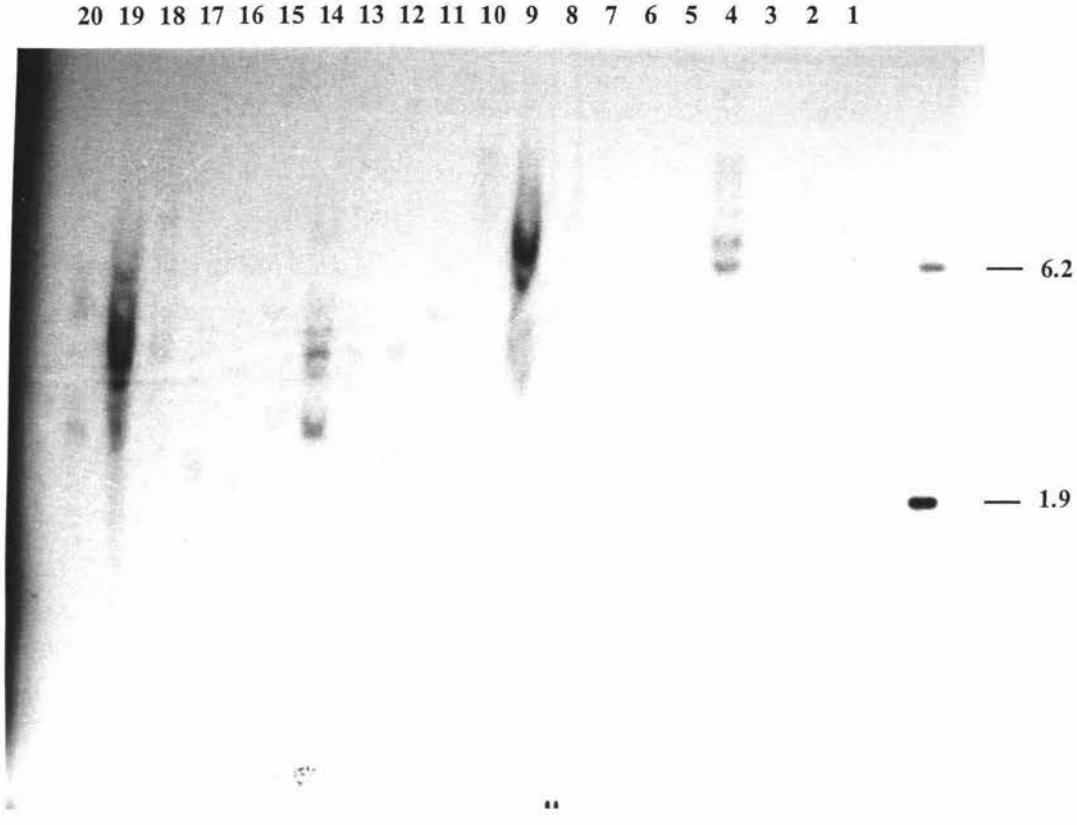


gel 10: *npt* II probed

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15



gel 11: *gus* probed



gel 12: *gus* probed